Temporal analysis of the least energetic events in pulsar data from observations with the high energy stereoscopic system

Aristeidis Noutsos

How to cite:

Aristeidis Noutsos (2006) Temporal analysis of the least energetic events in pulsar data from observations with the high energy stereoscopic system. Doctoral thesis, Durham University.

Temporal Analysis of the Least Energetic Events in Pulsar Data from Observations with the High Energy Stereoscopic System

— Volume 1 —

by

Aristeidis Noutsos

Submitted in conformity with the requirements
for the degree of Doctor of Philosophy

Department of Physics
University of Durham
South Road
Durham. UK.

Abstract

It has been more than 60 years since astronomers turned their attention towards the γ-ray window ($>$ 100 keV). Nowadays, γ-ray astronomy has won its place as a separate branch of astronomy in its own right. The present thesis introduces the reader to γ-ray observations in the $\sim$ 100 GeV–100 TeV energy window, but focuses, in particular, on the efforts to describe and detect the pulsed, Very High Energy (VHE) γ-ray emission from pulsars.

Pulsars are highly magnetised ($B \sim 10^{12}$ G), rapidly rotating ($P \sim 10^{-2}$ s) neutron stars. Periodic radio emission from pulsars has been detected in more than 1,500 cases, in contrast to their γ-ray signature which has been confirmed for only six of them and only up to a few GeV. There are many models in existence which attempt to reproduce the observed pulsed profiles and energy spectra in high energies (optical, X and γ rays). Nevertheless, two classes of models are the most popular: the Polar Cap and the Outer Gap models. Both predict spectral cut-offs at tens of GeV, which are consistent with previous upper limits in the VHE range.

The six most energetic pulsars have been detected with the EGRET (Energetic Gamma Ray Experiment Telescope) instrument on-board the *CGRO* (*Compton Gamma Ray Observatory*) satellite. Probing the universe at higher energies requires a different detection technique. The Imaging Atmospheric Technique (IACT) exploits the Earth's atmosphere with the use of large, ground-based reflectors that are very sensitive to Cherenkov light (300–600 nm). The latter is produced during electromagnetic particle cascades, triggered by the interaction of VHE γ rays with the top atmospheric layers. So far there has not been a confirmed pulsar detection using Cherenkov astronomy.

The High Energy Stereoscopic System (H.E.S.S.) in Namibia is an array of four telescopes, which is sensitive above 100 GeV. H.E.S.S. uses the IACT to reject the 10^3 times more abundant cosmic-ray events that suppress the γ-ray signal. The system is capable of stereoscopic observations of the same source with all four telescopes, which further eliminates background events. Despite the fact that imaging with H.E.S.S. is not effective below 100 GeV, lower energy events can still be recorded, along with a large portion of the background. The present thesis deals with the least energetic events ($\lesssim$ 100 GeV) detectable with H.E.S.S., where pulsar γ-ray emission is likely to be present. A very sensitive temporal analysis has been performed in order

to identify the potentially periodic events in the large background. The necessary procedures and parameters of the analysis are described in detail, prior to the results.

The author has analysed data from two γ-ray pulsars, the Crab and PSR B1706−44, which were seen with EGRET up to $\sim$ 20 GeV, as well as the binary radio pulsar PSR B1259−63, which has not been detected at high energies ($>$ 1 eV). The data were optimised for the lowest energies, and the lowest energy threshold achieved was 75 GeV (in the case of PSR B1706−44). In all cases studied, the author coded and applied a number of periodicity tests that check for significant deviations from random noise. The resulting probabilities were not significantly low to support signal presence. Based on the background levels in the data sets, the author derived upper limits on the integral and differential flux. These upper limits were consistent with the Polar Cap and Outer Gap scenarios, within statistical errors, but constrain the alternative model of a spectrum with a simple exponential cut-off in the case of PSR B1706−44. Despite the lack of detection, these results represent the lowest energies explored with H.E.S.S., yet.

Declaration

The material contained within this thesis has not previously been submitted for a degree at the University of Durham or any other university. The research reported within this thesis has been conducted by the author unless indicated otherwise.

Copyright Notice

Acknowledgements

I consider myself particularly fortunate not only to have been part of the University of Durham γ-ray astronomy group, but also to have been in the hands of two, very competent supervisors. I would like to thank Dr P. M. Chadwick and Dr S. M. Rayner, who aided my research with their long experience in the field. Further thanks should go to the University of Durham for covering the tuition fees during my two, final years. However, it would not have been possible to work in Durham without the help of my former mentor, Prof J.-H. Seiradakis, under the supervision of whom I completed my diploma thesis, and Prof S. Avgoloupis, whose astronomy lectures were proven very useful in my research. Also, I would like to acknowledge the help of the high-energy physics group of the Humboldt University of Berlin and especially Prof Dr T. Lohse, Dr A. K. Konopelko, F. Breitling and F. Schmidt.

In addition, I am obliged to my colleagues, who provided technical support when it was most needed. Especially, I would like to thank Dr K.-M. Aye for his constructive comments; Dr M. K. Daniel for his generosity in passing on his knowledge (and other useful material); and my colleague and compatriot Mr C. N. Hadjichristidis, whose presence made me feel at home. Had it not been for Mr I. J. Latham, the research process would have lacked its inherent fascination and in many of the occasions would have fallen short of my expectations. Dr R. Le Gallou's assistance and advice in many scientific and other issues played a central part in my research, and I would like to thank him for that. I would also like to thank Dr J. McKenny for the ever useful suggestions and Dr S. J. Nolan for providing his technical expertise, despite his busy schedule.

A special thanks must go to all residents of the UN house, who provided psychological support during those stressful hours, and to all our friends that believed in my effort. In particular Dr H. G. Svendsen, whose interest in philosophical debates helped me many times to escape from the long and painstaking process of research. Furthermore, I would like to thank Dr J. M. Weber, who has been a force of inspiration in many of my activities, and a great counsellor in personal matters. Also, I thank Ms G. M. Horrocks for the substantial support, especially when the writing of this document occupied most of my time.

Last but not least, a very special thanks goes to Karen Findlay for her continuous support, in every possible way, during the last couple of years, which has been a unique life experience for me.

Finally, I owe my gratitude to two loving parents, who provided both the necessary means and the mental boost ever since I was making my first steps in science … and in this world.

Dedication

I dedicate the whole of this work to my parents. Also, in recognition of her patience and understanding, I dedicate this thesis to Karen.

Σα βγείς στον πηγαιμό για την Ιϑάκη,
να εύχεσαι νά 'ναι μακρύς ο δρόμος,
γεμάτος περιπέτειες, γεμάτος γνώσεις.
Τους Λαιστρυγόνας και τους Κύκλωπας,
τον ϑυμωμένο Ποσειδώνα μη φοβάσαι,
τέτοια στο δρόμο σου ποτέ σου δε ϑα βρείς,
αν μεν' η σκέψις σου υψηλή, αν εκλεκτή
συγκίνησις το πνεύμα και το σώμα σου αγγίζει.
Τους Λαιστρυγόνας και τους Κύκλωπας,
τον άγριο Ποσειδώνα δε ϑα συναντήσεις,
αν δεν τους κουβανείς μες στην ψυχή σου,
αν η ψυχή σου δεν τους στήνει εμπρός σου.

Να εύχεσαι νά 'ναι μακρύς ο δρόμος.
Πολλά τα καλοκαιρινά πρωϊά να είναι
που με τι ευχαρίστησι, με τι χαρά
ϑα μπαίνεις σε λιμένας πρωτοειδωμένους·
να σταματήσεις σ' εμπορεία Φοινικικά,
και τες καλές πραγμάτειες ν' αποκτήσεις,
σεντέφια και κοράλλια, κεχριμπάρια κ' έβενους,
και ηδονικά μυρωδικά κάϑε λογής,
όσο μπορείς πιο άφϑονα ηδονικά μυρωδικά·
σε πόλεις Αιγυπτιακές πολλές να πας,
να μάϑεις απ' τους σπουδασμένους.

Πάντα στο νου σου νά 'χεις την Ιϑάκη.
Το φϑάσιμον εκεί ειν' ο προορισμός σου.
Αλλα μη βιάζεις το ταξίδι διόλου.
Καλλίτερα χρόνια πολλά να διαρκέσει·
και γέρος πια ν' αράξεις στο νησί,
πλούσιος με όσα κέρδισες στο δρόμο,
μη προσδοκώντας πλούτη να σε δώσει η Ιϑάκη.

Η Ιϑάκη σ' έδωσε τ' ωραίο ταξίδι.
Χωρίς αυτήν δε ϑάβγαινες στο δρόμο.
Άλλο δεν έχει να σε δώσει πια.

Κι αν πτωχική τη βρείς, η Ιϑάκη δε σε γέλασε.
Έτσι σοφός που έγινες, με τόση πείρα,
ήδη ϑα το κατάλαβες οι Ιϑάκες τι σημαίνουν.

—*Ιϑάκη*, Κωνσταντίνος Π. Καβάφης

As you set out for Ithaka
hope your road is a long one,
full of adventure, full of discovery.
Laistrygonians, Cyclops,
angry Poseidon — don't be afraid of them:
you'll never find things like that on your way
as long as you keep your thoughts raised high,
as long as a rare excitement
stirs your spirit and your body.
Laistrygonians, Cyclops,
wild Poseidon — you won't encounter them
unless you bring them along inside your soul,
unless your soul sets them up in front of you.

Hope your road is a long one.
May there be many summer mornings when,
with what pleasure, what joy,
you enter harbours you're seeing for the first time;
may you stop at Phoenician trading stations
to buy fine things,
mother of pearl and coral, amber and ebony,
sensual perfume of every kind —
as many sensual perfumes as you can;
and may you visit many Egyptian cities
to learn and go on learning from their scholars.

Keep Ithaka always in your mind.
Arriving there is what you're destined for.
But don't hurry the journey at all.
Better if it lasts for years,
so you're old by the time you reach the island,
wealthy with all you've gained on the way,
not expecting Ithaka to make you rich.

Ithaka gave you the marvellous journey.
Without her you wouldn't have set out.
She has nothing left to give you now.

And if you find her poor, Ithaka won't have fooled you.
Wise as you will have become, so full of experience,
you'll have understood by then what these Ithakas mean.

—*Ithaka*, C. P. Cavafy
(transl.: Edmund Keeley &
Philip Sherrard)

Contents

Chapter 1

The Foundations of γ-ray Astronomy

1.1 Introduction

After astronomy had been established in the radio and optical parts of the spectrum, interest turned to higher energies. Ground-based telescopes were not appropriate for the direct detection of high-energy photons: the Earth's atmosphere acts as a filter against certain bands of the electromagnetic (hereafter EM) spectrum (Fig. 1.1), which prevents lethal high-energy radiation from entering the biosphere. Instead, it was required to either take the detectors above the atmosphere, so that any attenuation could be avoided, or find the means for an indirect detection from the ground. The aim was to explore the non-thermal universe, where source emission is dominated by X-rays and γ rays. γ rays are of higher energy than X-rays, although that alone is not what distinguishes them from each other. Generally speaking, the X-ray band is defined in the energy range $0.1 \leq E \leq 100$ keV, and everything above is considered γ rays.

The soft X-ray band (0.1–30 keV) can only be explored with satellite observations. Above this range, the hard X-ray photons are energetic enough to penetrate a small fraction of the atmosphere ($\approx 0.3\%$), and one can use high-altitude balloons equipped with the appropriate detectors to detect them. Balloon experiments can also be used for the detection of γ rays above tens of MeV, i.e. in High Energies (HE). The photons in these energies can pair produce in the upper atmospheric layers and the $e^{-}-e^{+}$ pairs can be detected by the balloon's payload. However, satellite observations, with their ability to continuously track a source over longer periods, have been proven a lot more successful in HE. Furthermore, the recent development in instrumentation sensitivity for such observations will extend their energy coverage to as high as $\approx$ 100 GeV.

Above the HE range, indirect observations from the ground are possible. Probing the very-high-energy universe requires a slightly different approach to that used in traditional astronomical observations (i.e. optical and radio observations). At very high energies, the detectors have to be appropriately designed in order to detect the signatures of the energetic γ-ray photons. At such high energies, the light–matter interaction processes have to be exploited. Ironically, the principles of optical astronomy can still be used for the study of VHE γ rays, as we will see in chapter 4.

Very High Energy (VHE) γ rays, $\gtrsim$ 100 GeV[1], can trigger EM cascades in the atmosphere which can be detected by Cherenkov telescopes or water-pond arrays. The lack of source detections above 50 TeV with such detectors has indicated this energy as the natural upper limit for VHE γ-ray astronomy. Finally, at the very highest energies, fittingly named Ultra High Energies (UHE), the particle products of $\gtrsim$ 100-TeV γ rays are detectable with arrays of ground-based, particle or fluorescence detectors. A brief description of all of the above detectors is given in the following sections.

Table 1.1 presents the classification of high-energy observations according to energy. The reader should bear in mind that the limits for each energy range are not strict, and that they are the direct result of observability at the corresponding energies. Future development of the detectors could easily alter these ranges and potentially lead to an overlap between space-borne and ground-based detections, for example.

1.2 Space Detectors

1.2.1 Detection Techniques

A key issue with high-energy emission is that it is by a large percentage filtered by the atmosphere, and therefore direct detection is impossible below the Earth's atmosphere. Hence, the first approach to this fairly new branch of astronomy was

---

[1] The physical lower limit for Cherenkov emission in the atmosphere is a few GeV. However, the majority of Cherenkov detectors have thresholds exceeding 100 GeV; and this is what has defined the VHE region.

Detector type	Energy Range (Acronym)
Satellite	0.1 – 30 MeV (LE/ME)
Satellite/Balloon	30 MeV – 100 GeV (HE)
Cherenkov/Water pond	100 GeV – 100 TeV (VHE)
Particle/Fluorescence	> 100 TeV (UHE)

Table 1.1: (from [1]) Classification of the high-energy detector types according to energy. The acronyms LE, ME and HE correspond to Low, Medium and High Energy, respectively.

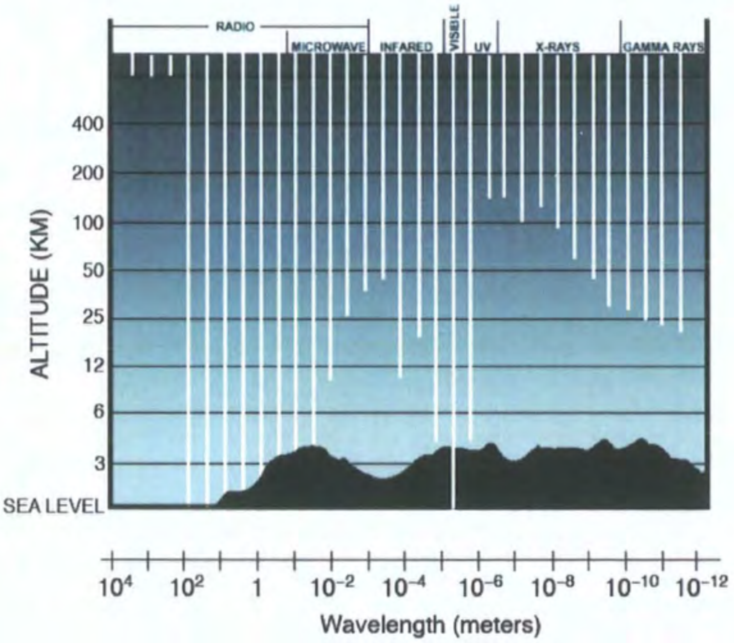


Figure 1.1: (from [2]) Transparency of the Earth's atmosphere to the electromagnetic spectrum.

to place the γ-ray detectors at a high altitude, so that they would not be confronted by such an obstacle. Initially, the devices were on-board high-altitude balloons that were designed to stay afloat at $\approx$ 40 km and usually for a couple of days, at best. For γ ray observations, floating at such high altitudes can be regarded as being 100% above the atmosphere. More recent experiments, like the Balloon-borne Experiment with a Super-conducting Spectrometer (BESS) and the Cosmic Ray Energetics and Mass experiment (CREAM), have the capability of continuous observation for up to 3 weeks or more [3],[4]. They can carry a 2,000-kg payload, and the detector's active area can be as large as 30 m^2.

However, as the satellite technology evolved, γ-ray detectors were integrated into satellite experiments, which provided the possibility for observing the various γ-ray sources for up to years. The importance of such long periods of observation can only be appreciated if we consider the very low flux received by a satellite detector. For example, EGRET's weakest γ-ray pulsar, PSR B1055$-$58, is detectable with that instrument above 100 MeV (when the source is within 10° of the telescope's axis) at the mere rate of 4 ph d^{-1} [5]. Therefore, for the statistics to be meaningful, long-term observations are required.

In order to respond to the different energy bands of γ-ray astronomy, the detector apparatuses are based on various physical principles. The starting threshold for γ-ray detection is defined by the annihilation energy of an electron–positron pair, which is 0.511 MeV. Satellite detectors aiming at the detection of γ rays up to about 3 MeV use either scintillation counters or, alternatively, the more expensive solid state detectors. The scintillation principle is based on the excitation of the scintillator's molecules — these can be inorganic (e.g. NaI) or organic (e.g. polystyrene) — by the energetic γ rays, and the re-emission of lower frequency light (usually optical or UV) from these molecules. The amount of re-emitted light can then be collected by a sensitive photomultiplier and translated into a measure of energy, by means of the scintillator's and photomultiplier's properties (i.e. quantum efficiency, gain, etc.). A clear disadvantage of such detectors is that they require a large scintillating mass in order to absorb the entire energy of the γ-ray photons via the multiple excitation and re-emission process. One useful advantage, though, is their good temporal resolution that allows the accurate estimation of event arrival times.

Satellite γ-ray telescopes carrying solid state detectors operate on roughly the same principle. The difference is that when a solid state detector is struck by an ionising γ ray, the detector's atoms are ionised rather than excited, and therefore electron–hole pairs are created. By applying an electric field across the semiconductor, the liberated charges (electron and hole) are forced to move in opposite directions and therefore create a measurable current. Knowledge of the semiconductor's properties allows for translation of such a current into photon signal. Unfortunately, the required effective area ($\sim 10^3$ cm^2) for such experiments pushes the designers towards large amounts of semiconducting material (e.g. Ge–Li), which can be very expensive. Nevertheless, the gain from such devices is in terms of energy resolution: for example, values of $E/\Delta E \approx 500$ can be achieved [6].

At energies of a few MeV, a typical system involves Compton scattering of the γ rays by silicon detectors, for example. The basic geometrical configuration of a Compton detector is depicted by Fig. 1.2.

As the γ ray photon enters the detector's mass, there is a high probability, at those energies, of interaction (collision) between the photon and one of the electrons of the lattice. In the figure, the photon interacts first with *detector 1*; then the detector's electron immediately absorbs part of the photon's energy and shoots off into one direction, while the photon is scattered in a different one. Compton's formula relates the change in the photon's direction, by an angle Φ, with the photon's energy, E_γ, and the energy lost in *detector 1* through the photon–electron collision, E_1. Of course, in order to measure E_γ, there has to be enough scattering material to absorb all the photon's energy. Therefore, the Compton telescope uses at least two detector plates some distance apart — in the present figure, detectors *1* and *2* — which serve two purposes: one, to determine the γ ray's total energy, which equals $E_\gamma = E_1 + E_2$ (assuming that *detector 2* has absorbed all of the remaining photon's energy, E_2); and two, to determine the photon's direction after it has been scattered. The scattered photon's direction can be simply reconstructed by joining the two interaction points E$_1$ and E$_2$ with a straight line (dashed, red line in Fig. 1.2). Having this information, one can calculate the angle Φ using Compton's formula, which in our case translates to

$$\cos(\Phi \pm \Delta\Phi) = 1 - m_e c^2 \left(\frac{1}{E_\gamma - E_1} - \frac{1}{E_1} \right) \tag{1.1}$$

where $m_e c^2$ is the rest energy of the electron, E_γ is the total energy of the incident γ-ray photon (in the electron's reference frame), and E_1 is the energy absorbed by *detector 1*; $\Delta\Phi$ is the uncertainty in the calculation of Φ, which arises from the systematic uncertainties involved in the measurement of E_1 and E_2, as well as the direction of the scattered photon.

The determination of Φ restricts the angular distance between the axis E_1E_2 and the direction from which the incident γ-ray has come; but the azimuthal direction is not defined. Therefore, the geometrical region which contains all the possible incident directions is a cone of half-angle Φ, with axis the reconstructed direction of the scattered photon (E_1E_2). Allowing for the uncertainty $\Delta\Phi$ and projecting this cone on the plane of the sky gives us a ring of angular radius $\Phi \pm \Delta\Phi$ that covers all the possible directions.

A Compton telescope can also operate by reconstructing the direction of the primary photon via the projection not of the scattered photon, but of the scattered electron. Although this variation of the detector can have a higher yield, since the absorption of the electron is more likely than that of the energetic secondary photon, the low-energy electrons can suffer multiple scattering while passing through the detector, which results in less accurate determination of the energy and direction of the primary.

At energies above 10 MeV, there is another interaction between light and matter that can be used to our benefit. Pair production becomes more dominant than Compton scattering in this energy regime, and moreover its products define the direction of the primary γ-ray in a less ambiguous way. The basic principle of pair production is discussed later in this thesis, but in simple terms it is the product of the interaction between a high-energy photon and the Coulomb field of a nucleus. This interaction leads to the transformation of the photon into an $e^{-}-e^{+}$ pair with kinetic energy equal to the difference between the rest energy of the pair and the energy of the photon. The process can only occur at energies higher than 1.022 MeV, but its cross section becomes important, for experimental purposes, at higher energies.

The basic instrument that is used for the detection of γ rays via pair production is the spark chamber. It consists of parallel metal plates placed in a gas chamber that is usually filled with an inert gas, like argon or neon. The plates are alternately connected to the ground and to a high-voltage source, but the high voltage is not always switched on, so that lossy ionisation of the gas by intruding charged particles can be avoided. Instead, it is required that a co-operative detector has been triggered by the first interactions of γ rays with the metal plates, in order for the voltage switch to be activated. This fact alone is enough to raise the energy threshold of the technique to about 10 MeV. After the spark chamber has been activated, each γ ray

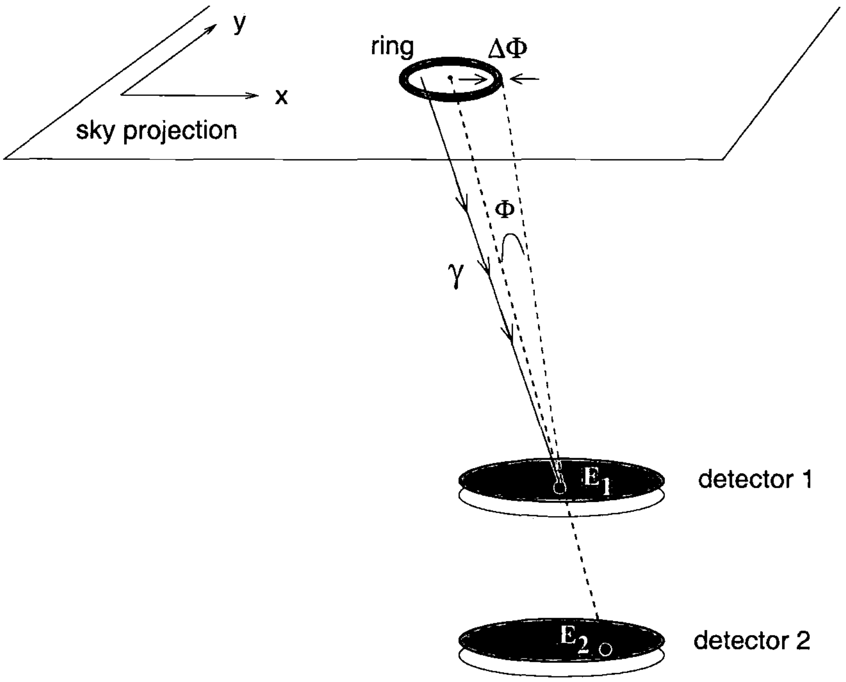


Figure 1.2: (from [7]) The direction of the initial γ ray detected by a Compton telescope can be constrained by Compton's law. This requires the photon's total energy, which is calculated by adding the energy expenditure on *detector 1* to the energy of the scattered photon, absorbed by *detector 2*. The calculated angle, Φ, can then be added to the projected direction (the dashed, red line defined by the impact points E1 and E2), in order to project a circle on the sky which shows from where the initial γ ray might have come. The width of the projection, $\Delta\Phi$, corresponds to the uncertainty in determining Φ. As a result, the projection now becomes a ring section covering all the probable arrival directions.

that is converted into an $e^{-}-e^{+}$ pair is detected as a trail of sparks which is formed by the strongly accelerated ions in the high electric potential. Photography or some other electronic method can be used to record the sparks and translate them into γ-ray counts.

The angular resolution in such an instrument is superior to the Compton telescope's, due to the fact that the $e^{-}-e^{+}$ pairs follow a path almost parallel to the primary photon's. These particles — which are highly energetic — do not suffer from multiple scattering, and the subsequent positions of the sparks form a direction that matches closely the primary photon's. Otherwise, the uncertainty in the determination of the photon arrival directions depends only on the chamber's characteristics [8].

1.2.2 Early Observations

In May 1961, the first satellite capable of detecting γ-rays was launched. In its 7-month operation, *Explorer XI* managed to detect 1012 γ-ray photons. The fraction of these events that were regarded as coming from extraterrestrial sources was 31, whereas the rest were produced from Earth's direction, from interactions of high-energy particles with the atmosphere. Using this limited sample, the inferred upper limit on the extraterrestrial γ-ray intensity was $J_{\mathrm{ul}}(> 100\ \mathrm{MeV}) = 3.3 \times 10^{-4}\ \mathrm{cm}^{-2}\ \mathrm{s}^{-1}\ \mathrm{sr}^{-1}$ [9].

After *Explorer XI* opened the way to satellite observations of γ rays, many other experiments followed. The first satellite to be sent to space for a specific γ-ray mission, *SAS 2*, was launched in 1972 and was followed shortly after by *COS-B*. Both experiments provided an insight into γ-ray astronomy. More specifically, *SAS 2* verified the diffuse γ-ray emission that had been picked up by *Explorer XI*, and *COS-B* provided a detailed map of the Galactic γ-ray emission and discovered 25 discrete γ-ray sources.

Amongst the γ-ray sources discovered with *COS-B* are the Crab and Vela pulsars, as well as the quasar 3C 273. The first two objects, i.e. young pulsars surrounded by expanding supernova shells, are very well known in every sort of astronomical research, from radio waves to γ rays. *COS-B* managed to detect a highly variable, pulsed γ-ray profile from the Crab pulsar at energies up to 3 GeV, as opposed to the relatively steady profile of the Vela pulsar [10]. PSR J0835−4510's profile, as was seen with *COS-B*, is shown in Fig. 1.3. Apart from those two which are considered standard candles for γ-ray astronomy, no other pulsars were seen to emit pulsed radiation.

Furthermore, the *COS-B* mission detected extragalactic γ rays from the object 2CG 289+64, which was later identified as the quasi-stellar object 3C 273. The high-energy emission from this object has been extensively studied with subsequent satellite experiments, like EGRET and COMPTEL. An investigation of the multi-wavelength energy density spectrum shows that 3C 273 peaks at energies of a few MeV: i.e. in hard X-rays or, equivalently, soft γ rays (see Fig. 1.4). An interesting feature of 3C 273's high-energy emission is the temporal variability of the flux on a monthly time-scale. The observed flux above 100 MeV (i.e. $dN/dE(E > 100\,\mathrm{MeV}) \propto E^{-\nu}$) is seen to vary by a factor of 3 between the object's more energetic "high" state and its "quiet" state. This flux increase is due to the hardening of the energy spectrum by $\Delta\nu \approx -1$ during the high state (see Fig. 1.5) [11].

1.2.3 Modern Instruments

Perhaps the most successful γ-ray space instrument was EGRET on-board the *CGRO* satellite [13]. EGRET utilised a spark chamber to detect γ rays above 20 MeV via pair production. Its most important contribution emerged from the 1991–1995 all-sky survey. The results from this survey were used to compile the EGRET catalogue of γ-ray sources, which is composed of 271 objects — most of them unidentified due to the feeble spatial resolution [14]. Amongst them are pulsars, SNRs, AGN, etc. Only six of the γ-ray sources detected in the survey were confidently associated with known radio pulsars (see Table 1.2). Further pulsar associations required contemporaneous observations of the unidentified sources with EGRET and radio/X-ray experiments: this is because the detected γ-ray fluxes were very low and required an accurate ephemeris, in order to fold the few detected events accurately into lightcurves. However, EGRET's operation ceased in 2000 and the potential associations had to be based on different criteria [15]:

- The HE spectra of the unidentified sources were compared with those of the confirmed γ-ray pulsars: sources with spectral indices that were close to the already known sample ($\nu = 1.42$–2.1) were considered as stronger candidates than sources with much softer (say $\nu = 2.6$) spectral indices.

- Probabilistic studies were performed to calculate the probability that the number of considered associations between EGRET sources and pulsars from the

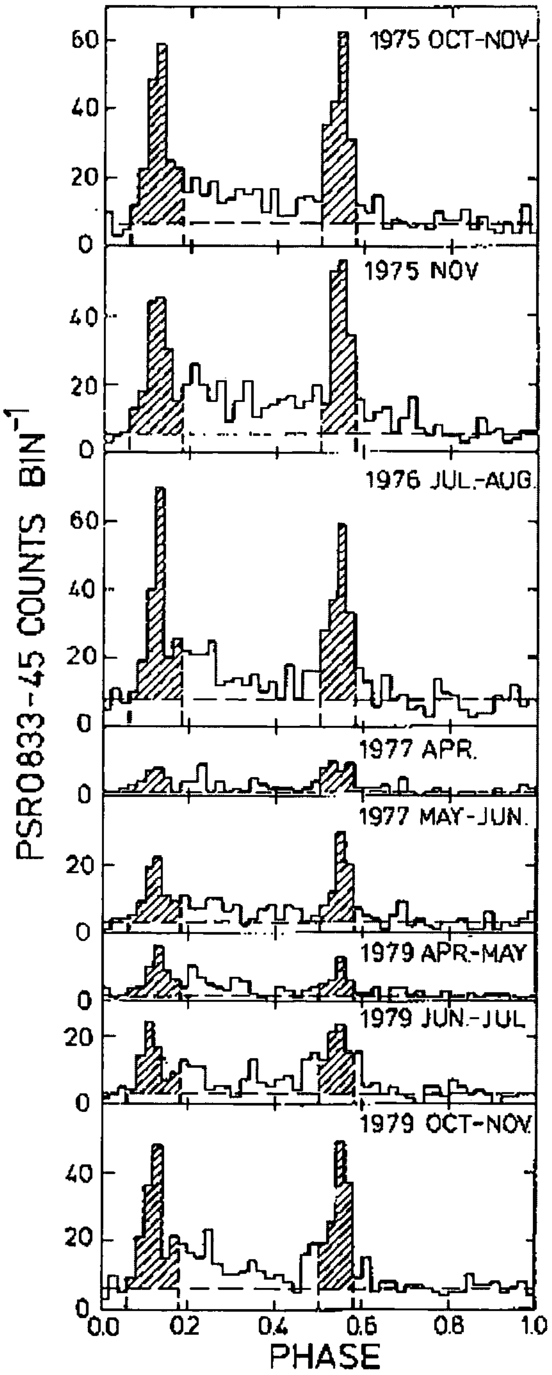


Figure 1.3: (from [12]) The Vela pulsar's (PSR J0835−4510) lightcurves for different periods of *COS-B* observations.

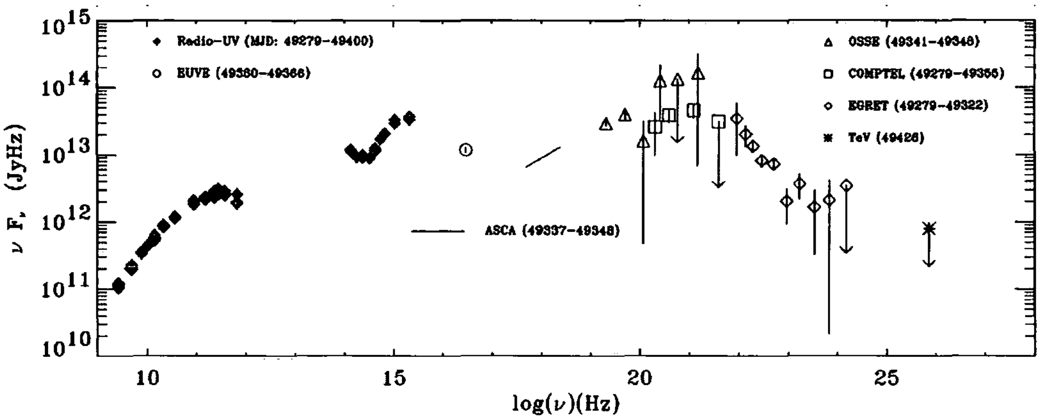


Figure 1.4: (from [11]) The multi-wavelength spectrum of QSO 3C 273 from various campaigns, including observations with the *CGRO* (*Compton Gamma Ray Observatory*) instruments, EGRET ($E > 100$ MeV) and COMPTEL ($0.75 < E < 30$ MeV).

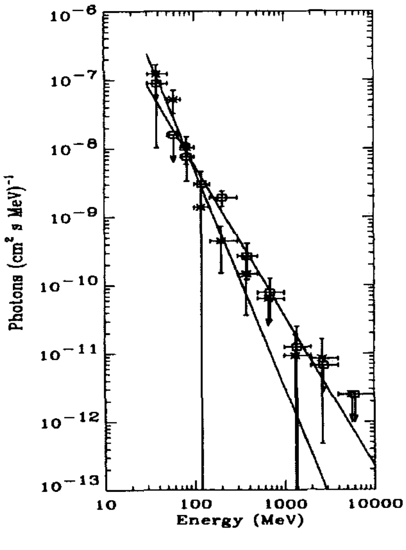


Figure 1.5: (from [11]) The high-energy (> 100 MeV) spectra of 3C 273 during its "high" (squares) and "quiet" (asterisks) state, as were observed with the EGRET instrument on-board *CGRO*. The spectral index of the power-law fit to the high state data is $\nu = 2.20$, whereas the quiet state is described by a power law with $\nu = 3.20$.

Parkes pulsar survey would occur by chance: e.g. for a sample of 8 independent associations, a 2% chance has been reported.

- Based on EGRET's fluxes and the distances to EGRET's sources, the γ-ray flux as a percentage of the spin-down flux (γ-ray conversion efficiency; see section 3.1.3) was calculated and compared with that of the known sample. Again, values of $\lesssim 20\%$ were favoured against higher, more unphysical ones.
- Finally, sources which showed variable (time-averaged) emission were disfavoured as being pulsars, as this is not expected from the latter objects.

Aboard the same satellite, a different instrument, which used a Compton scatterer for source localisation and spectroscopy, was the COMPTEL. COMPTEL had been used for the study of various Galactic and extragalactic high-energy sources in the 0.75–30 MeV energy range, including pulsars, plerions, GRBs, black-hole candidates, etc. [16],[17].

One of the modern satellite missions, which utilises solid state detectors for γ-ray imaging and spectroscopy, is *INTEGRAL*. Its recent achievements include the detection of the closest GRB yet, GRB 031203, and the resolution of the Galactic γ-ray emission, which was previously thought to be diffuse, into distinct point sources [18],[19].

Another satellite mission which began observations in 2004 is *SWIFT*. It employs the Burst Alert Monitor (BAT), i.e. a solid-state, Cadmium–Zinc–Telluride (CZT) detector whose main purpose is the localisation of GRBs — which is also the main target of *SWIFT*'s mission [20].

EGRET's successor, *GLAST* (Gamma-ray Large-Area Space Telescope), is another of NASA's mission, whose launch is anticipated in 2007 [21]. The telescope's primary instrument, LAT (Large Area Telescope), is an imaging detector that comprises four tower modules embedded in a plastic scintillator shield used for anti-coincidence trigger. Each tower contains an array of parallel sheets of ^{28}Si-strip detectors (SSDs) interleaved with ^{184}W sheets. The former are used as a tracker: i.e. to track the path of the charged particles created from the pair production of the incident γ rays on the tungsten sheets (the converters). Every interaction with the silicon sheets is recorded via voltage pulses which reveal the particles' trace on the silicon planes. The 3-dimensional trajectory is then reconstructed using the subsequent traces as the particles move across the detector.

The LAT instrument on-board *GLAST* will provide significantly higher angular resolution compared to EGRET: it is expected to be able to resolve sources to within < 0.4 arcmin at 10 GeV. Moreover, it is hoped that *GLAST* will extend EGRET's energy range to ≈ 300 GeV and, thus, overlap with VHE observations (see section 7.2).

1.3 Atmospheric Detectors

1.3.1 Properties of the Detected Radiation

Sources of γ-ray emission produce fewer photons in VHE energies than in HE. For example, the unpulsed γ-ray flux from the Crab nebula in the 30–100 GeV energy range, as was seen with EGRET, ranges from 10^{-7} cm^{-2} s^{-1} MeV^{-1} to less than 10^{-13} cm^{-2} s^{-1} MeV^{-1} (see Table 1.3). Hence, as one moves into the GeV–TeV energy range, the photons become really scarce. As a consequence, in order to collect enough photons for a statistically useful sample, either the *exposure time*, which is the amount of on-source time, or the collection area, which is proportional to the detector's size, has to be increased. However, as happens with many of space missions, operational time can be very expensive, and the launch costs for a large space detector are enormous. Using a simple approximation for the total number of photons, N_γ, collected during exposure time T with a satellite of collection area A, one can calculate the required *exposure* (i.e. $A \times T$ cm^2 s) that would result in the same number of photons at 100 MeV and at 100 GeV. From Table 1.3, we have that the flux varies by about 5 orders of magnitude between those two energies, which means that under the same amount of exposure time EGRET's collection area ($A = 1,500$ cm^2) would have to be $\sim 15,000$ m^2: i.e. the size of a football stadium! Therefore, it is clear that typical space-borne experiments are inefficient at VHE observations.

Luckily, at VHE energies, all γ rays penetrating the top layers of the atmosphere pair-produce in the presence of atmospheric nuclei. The products carry the energy and momentum of the incident γ rays and give rise to further high-energy radiation via Bremsstrahlung. These secondary γ-ray photons pair-produce again, and the EM cascade develops in this way until the secondary products become thermal. A more detailed presentation of the development of such cascades in the atmosphere, called extensive air showers (EAS), is given in section 4.1.1.

During the initial stages of an EAS development, a large percentage of the sec-

Pulsar	P(ms)	$-\log\dot{P}$	$\log(B_s/\mathrm{G})$	$\log(\tau/\mathrm{y})$	d(kpc)	$\log\left[\frac{\dot{E}(\mathrm{ergs\cdot s^{-1}})}{d^2(\mathrm{kpc})}\right]$	Reference
Crab	33	12.4	12.6	3.1	2.0	−4.9	[22]
Vela	89	12.9	12.5	4.1	0.5	−5.5	[23]
B1706-44	102	13.0	12.5	4.2	1.8	−7.0	[24]
B1951+32	39	14.2	11.7	5.0	2.5	−7.2	[25]
Geminga	237	14.0	12.2	5.5	0.2	−7.3	[26]
B1055-52	197	14.2	12.0	5.7	1.5	−8.9	[5]

Table 1.2: (from [27]) The properties of 5 EGRET pulsars included in the 3^{rd} EGRET catalogue, plus PSR B1951+32 which was also detected by EGRET after extended pulsed analysis.

Energy range (GeV)	dN_γ/dE (cm^{-2} s^{-1} MeV^{-1})
0.03–0.05	1.5×10^{-7}
0.1–0.15	4.5×10^{-9}
0.5–1	2.8×10^{-11}
4–10	2.1×10^{-12}
30–100	$< 5.6 \times 10^{-14}$

Table 1.3: (from [28]) EGRET unpulsed fluxes from the Crab nebula, for different γ-ray energy bands.

ondary particles traverses the atmosphere at superluminal velocities. Then, due to a well-known effect called Atmospheric Cherenkov Effect, copious amounts of optical Cherenkov photons are generated along the particle paths and propagate in the shower's forward direction. A brief explanation of the physics involved in the Cherenkov light production in the atmosphere can be found in chapter 4.1.

Unluckily, the Atmospheric Cherenkov Effect is not a unique property of γ-ray-initiated EAS and it also occurs when the vastly more abundant cosmic rays (hadrons) enter the atmosphere. One can infer at this point that the key in doing γ-ray astronomy by using Cherenkov light from EAS lies in the ability to discriminate between γ-ray- and cosmic-ray-induced Cherenkov light. This task can be extremely difficult. The various techniques developed so far for this purpose exploit the directional information of the EAS, as well as their geometrical properties. The most successful yet, however, has proven to be stereoscopic imaging: a technique which is explained in section 4.2.6.

1.3.2 Characteristics of the Detectors

Cherenkov optical emission covers a very broad range of frequencies, from UV to IR, but the vast majority of Cherenkov photons are emitted in the blue and near-UV region of the spectrum, as shown in Fig. 1.6. For example, 1-TeV γ rays produce Cherenkov emission mostly in the 300–400 nm range. This light can be detected at ground level by Atmospheric Cherenkov Telescopes. In principle, these telescopes consist of a large reflective dish which can focus light on very sensitive photomultiplier tubes (PMTs). These PMTs, often called *pixels*, convert light into electrical signal. The total Cherenkov light across the night sky is $\sim 10^4$ times weaker than the total starlight, and so traditional methods of long exposure used in optical astron-

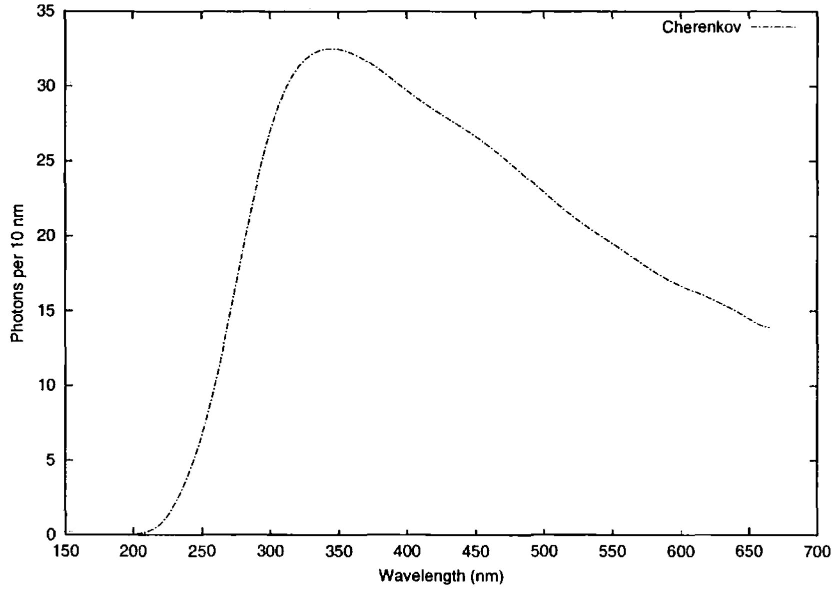


Figure 1.6: The Cherenkov differential spectrum corresponding to a 1-TeV γ-ray shower.

omy, for example, cannot be used in Cherenkov astronomy [29]. However, Cherenkov light generated from EAS lasts only for a few nanoseconds and is strongly beamed in the shower's forward direction. For that short amount of time and in the relevant direction, a Cherenkov flash outshines the Night Sky Background (NSB) light. Cherenkov telescopes are optimised for the registration of nanosecond-wide events that come from the direction of a source which is expected to emit.

Nevertheless, Cherenkov radiation can still be faint when it is induced by low-energy particles or low-energy γ rays. In addition, atmospheric attenuation can render even an energetic EAS undetectable if the Cherenkov light arrives at the detector after crossing a large distance through the atmosphere. This is typically the case for observations at large zenith angles (Z.A.s): i.e. for Z.A. $> 40^\circ$. Furthermore, the Cherenkov light pools spread out as they propagate towards low altitudes. Hence, it is optimal to have telescope dishes not only as large as possible but also at a high altitude, so that they can collect a large percentage of the Cherenkov light pool. In addition, the electronics in such telescopes have to be fast and sensitive in order to be able to respond to the short Cherenkov flashes. Finally, because the detectors are very sensitive, it is imperative to make Cherenkov observations during moonless nights and away from sources of light pollution (e.g. cities and airports).

1.3.3 Early Observations

There have been many attempts at the detection of γ-ray-induced, atmospheric Cherenkov emission, and there are many experiments currently operating around the globe for that purpose. The earliest experiment was performed in the U.K. by Galbraith and Jelley, at the Atomic Energy Research Establishment (AERE) in Harwell. Their 'telescope' had a 25-cm diameter, parabolic army-surplus mirror, mounted at the bottom of a dustbin. The Cherenkov light was being reflected on the mirror and detected with a 5-cm diameter PMT: the latest in light-detection technology, at the time [30]. Those early efforts were followed by another famous attempt, on the coast of the Black Sea. It was a group from the Lebedev institute, which observed during the period 1960–1964 in the Crimea using 12 1.5-m reflectors [31]. The mirrors were mounted on railway cars in groups of three, and the observation technique was similar to that of AERE.

Although those early experiments did not manage to detect any γ-ray sources at a significant level, many innovations were used: for example the fourfold coincidence technique, which required the detection of an event within a set amount of time by four neighbouring PMTs, in order for it to be regarded as a Cherenkov event. Both experiments mentioned above had very high energy thresholds (10 and 5 TeV), and their discrimination techniques were still primitive. However, they pioneered this branch of γ-ray astronomy and led the way for modern Air Cherenkov telescopes.

In 1968, the Whipple collaboration initiated an effort on Mt. Hopkins, Arizona (110°.8 W, 31°.6 N), at an altitude of 2.3 km. The Whipple telescope is a 10-m, Davies–Cotton reflector that consists of 248 tessellated mirrors. It focuses Cherenkov light on a high-resolution camera which provides it with a 4° field of view (FoV). One of Whipple's successes has been the clear detection of the Crab nebula above 250 GeV and its γ-ray spectrum reconstruction [32].

Another pioneering group that commenced VHE observations at the beginning of the 1980s was the U.K. group from the University of Durham. Their first two experiments, Mark 1 and 2, were situated in Dugway, U.S.A. (112°.9 W, 40°.2 N). Mark 1 was a configuration of four telescopes arranged at the centre and apices of an equilateral triangle with 100-m-long sides. Each telescope carried three search-light mirrors, focusing Cherenkov light onto single PMTs. Although the reflectivity of the system was not ideal, the arrival times of the individual events at the different telescopes were efficiently used to reconstruct the shower direction and reject off-source events. For that purpose, the Durham group pioneered in this branch of astronomy the threefold coincidence technique. The resultant energy threshold of those first telescopes was $\approx$ 1 TeV.

Two important contributions with this instrument were the discovery of pulsed γ-ray emission above 1 TeV from the Crab pulsar and the detection of high-energy emission from Cyg X-3. However, the claimed discovery of periodic TeV emission from the Crab pulsar, at the significance level of 4 σ, with this system, is yet to be confirmed by modern experiments.

1.3.4 Modern Instruments

Durham

The Durham group continued their γ-ray observations in Dugway until 1985, with the improved Mark 2 system. It had a three-mirror design similar to Mark 1, but each reflector consisted of 7, tessellated round facets. After the first two Mark telescopes, the Durham group turned to southern-sky observations from Narrabri, New South Wales, Australia (149°.8 E, 30°.3 S), with Marks 3, 4, 5 and 6. All of these telescopes were designed based on the original three-mirror design of the Durham telescopes. Amongst others, the Mark 3 and 4 telescopes were used for observations of several X-ray binary systems. In particular, the Mark 4 telescope was deployed at La Palma on the Canary Islands (17.89° W, 28.76° N) for specific northern-hemisphere observations of Cyg X-3, and was later transported to Narrabri.

It is worth noting that by the time the Durham Mark 5 telescope became operational in 1992, γ-ray astronomy had taken steps forward by introducing the Imaging Atmospheric Cherenkov Technique (see section 4.2.4), which was encompassed in the Mark 5 and 6 designs. This provided more efficient rejection of cosmic-ray-induced events and, in the case of the Mark 6 telescope, lowered the energy threshold to $\approx$ 300 GeV and made it comparable to the Whipple telescope's [33]. Amongst the contributions of the Mark 6 telescope, which ceased its operation in March 2000, was the confirmation of PSR B1706−44 as a high-energy emitter of unpulsed γ rays above 300 GeV [34].

Whipple/VERITAS

A follow-up experiment based on the original Whipple design is being built and operated by the VERITAS (Very Energetic Radiation Imaging Telescope Array System) collaboration [35]. The first phase of the experiment (Phase-I) will be composed of an array of four 12-m γ-ray reflectors, each one substantially improved from the Whipple design. The full stereoscopic array is scheduled to be complete by the end of 2006.

In January 2005, the first telescope of VERITAS was installed at Horseshoe Canyon on Kitt Peak, Arizona, at an altitude of 1800 m. Its 12-m-diameter dish carries 350 individual mirror facets providing a total reflective area of $\approx$ 110 m^2. Recent optimisations to the telescope's alignment and drive system have provided an angular resolution as low as $\approx 0.06°$ and maximum slewing speeds as high as 1°/s, respectively. The telescope's camera comprises 499 PMTs which provide a 3.5° FoV. Upon completion of Phase-I, all the telescopes will share the same characteristics as the existing one [36].

Once the first phase of VERITAS is complete, the system will be operated in stereoscopic mode, which allows for simultaneous observation, with all telescopes, of the individual Cherenkov showers. The first observations of the Crab nebula showed that a single VERITAS telescope is capable of reaching a threshold of $\sim$ 150 GeV (before event-selection cuts were applied to the data). Monte Carlo simulations of the 4-telescope array have shown that this threshold will be reduced to $\sim$ 110 GeV in stereoscopic observations, and the overall sensitivity of the array will be increased significantly [37].

More recently, in Autumn 2005, the VERITAS collaboration installed a second, identical telescope at the selected site, 85 m away from the first telescope. Hence, more sensitive, stereoscopic observations with the 2-telescope system are expected in the near future.

H.E.G.R.A.

The first experiment to utilise stereoscopy in their observations was that built by the H.E.G.R.A. (High Energy Gamma Ray Astronomy) collaboration. The observatory was located at La Palma on the Canary Islands (17.89° W, 28.76° N), at an altitude of 2.2 km, which is an optimal location because of the excellent atmospheric clarity and also because of the high altitude. The initial stage of the H.E.G.R.A. experiment involved a single, 5-m^2 reflector with an energy threshold of 1.5 TeV, which achieved the detection of the Crab nebula at a 6-σ level [38]. Further development of the experiment resulted in the addition of 5 telescopes which were able for the first time to function co-operatively, thus forming the first stereoscopic Cherenkov telescope array. As well as increasing the effective area of the system, the additional telescopes were capable of better γ/hadron discrimination, which pushed the energy threshold of the array down to $\approx$ 500 GeV. This led to the first detection of a shell-type supernova remnant in the northern sky, at sub-TeV energies [39]. The success of the first stereoscopic observations had a large impact on follow-up experiments, which readily welcomed the technique. A more detailed description of the operation of a Cherenkov telescope in stereo mode is given in section 4.2.4.

CANGAROO

Another big project, contemporary with H.E.G.R.A., was inaugurated by the CANGAROO (Collaboration between Australia and Nippon for a Gamma Ray Observatory in the Outback) in Australia (136°.7 E, 31° S). The collaboration's first telescope, CANGAROO-I, had a single, 3.8-m, composite reflector and a camera with a 2.7° FoV. Above the estimated energy threshold of $\sim$ 500 GeV, the Australian–Japanese collaboration managed to marginally detect the Crab nebula and claim a detection of non-periodic γ-ray emission from PSR B1706−44 [40]. Their latest upgrades to the experiment have been the CANGAROO-II and -III projects. CANGAROO-II is a 10-m parabolic dish that has 114, tessellated, spherical facets, each one 80 cm in diameter. Its camera consists of 552 pixels (i.e. PMTs), which offers improved Cherenkov imaging compared to the original design. The collaboration's latest experiment, CANGAROO-III, became operational in March 2004 [41]. It is a stereoscopic array of four 10-m telescopes — incorporating the CANGAROO-II telescope — whose aim is to explore the sub-TeV γ-ray sky. Each of the four telescopes carries a parabolic, segmented reflector with 114 small, spherical mirrors. The total reflective area of CANGAROO-III is 54 m^2. Three of the telescopes reflect the incident Cherenkov light onto an imaging camera which has 427 pixels and a 4° FoV, whereas the oldest of the four telescopes, CANGAROO-II, has a significantly narrower FoV (2.7°) but a higher-resolution camera (552 pixels). Due the narrower FoV of the latter telescope, but also due to the fact the it has a lower efficiency in stereo observations, its operation has ceased [42].

As a stereoscopic array, CANAGROO-III achieves an angular resolution of $\approx 0.3°$ and an energy resolution of 35% for observations up to 20° zenith angle. Currently, the estimated energy threshold of the array is $\approx$ 600 GeV [43].

H.E.S.S.

In June 2002, the first telescope of the High Energy Stereoscopic System (H.E.S.S.) became operational, and by the end of 2003 all 4 Cherenkov telescopes of the stereoscopic array were brought to completion. The H.E.S.S. location was chosen carefully in order to conform with the requirements of γ-ray observations, such as the high altitude, the dark skies and the good, on average, weather. A decision was made to place H.E.S.S. on the Khomas Highland of Namibia (16°.5 W, 23°.2 S), which is a fairly isolated place in the southern hemisphere, at an altitude of 1.8 km. Each of the 4 H.E.S.S. telescopes consists of 382 spherical mirrors arranged in a Davies–Cotton fashion, on a 13-m diameter dish. The individual mirrors focus the Cherenkov light on a high-resolution camera consisting of 960 pixels that cover a wide 5° FoV.

Upon completion, H.E.S.S. became the most advanced instrument of its kind, with an energy threshold of 100 GeV (for spectroscopy), but also the sensitivity to detect events below that energy [44],[45]. Its pointing accuracy, which is more than 3 times better than that of other contemporary experiments [46], allows H.E.S.S. to chase point sources, e.g. pulsars. But also H.E.S.S. can use its superior angular resolution (source localisation to within $\sim$ 20 arcsec) to map the morphology of extended γ-ray sources, e.g. SNRs.

Hence, the challenge for H.E.S.S. was the detection of faint signals, which would open the way for the study of a vast sample of γ-ray sources that were too faint to be detected with previous experiments (see Fig. 1.7). Future expansion of the system has been planned with the construction of a fifth, very large telescope, positioned at the centre of the current configuration. With this additional telescope, the system will boast an impressive, for its time, total reflective area of 1,032 m^2, which will provide detectability of γ rays down to 20 GeV [47]. An detailed description of the H.E.S.S. array is given in section 4.3.

MAGIC

2004 saw the emergence of another big γ-ray project. Built on the same site as the discontinued H.E.G.R.A. telescopes, MAGIC is currently the largest, single Air Cherenkov reflector, with a reflective surface amounting to 234 m^2. Its location (2,200 m a.s.l.) combined with the telescope's large reflective area is expected to give MAGIC the lowest threshold in its class: a threshold energy as low as 30 GeV has been projected. The telescope construction was based on many innovations:

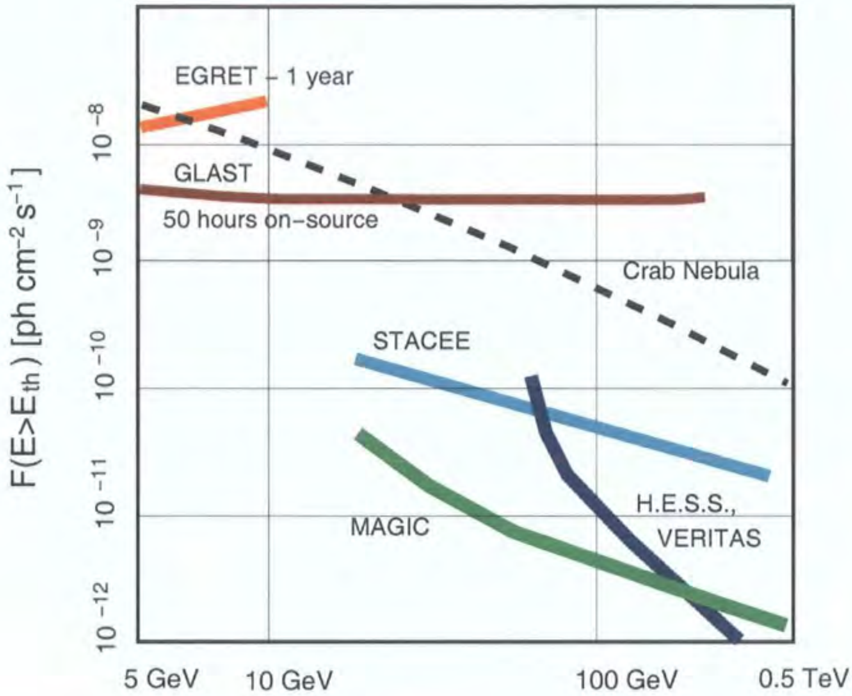


Figure 1.7: (from [51]) The expected sensitivity of MAGIC compared with that of other γ-ray experiments. The sensitivity ranges are drawn based on the minimum flux required for a 5-σ detection, after 50 h of exposure time.

e.g. a carbon-fibre dish frame, which allows for fast slewing (< 20 s average tracking time); a super-efficient camera with PMTs of 10–20% increased Quantum Efficiency (QE), relative to that of other experiments; and the use of laser beams for on-the-fly, individual mirror calibration [48].

The MAGIC collaboration is planning to build at their site a second, similar telescope, called MAGIC-II, which will provide stereoscopic capabilities to the system and lower even further the energy threshold by a factor ~ 2. Its construction is scheduled to finish sometime in 2007 [49].

Observations with the first telescope (MAGIC) commenced in August 2004. In the first months of operation, the instrument underwent the obligatory fine-tuning with the help of the strongest γ-ray sources, like the Crab nebula and Mkn 421. MAGIC is expected to be more sensitive than its contemporaries in the 30–100 GeV range (see Fig. 1.7), and therefore it is scheduled to observe sources of emission at the tail-end of the VHE spectrum, like the EGRET γ-ray pulsars [50].

STACEE/CELESTE

A quite different approach to ground-based γ-ray observations of air showers is the use of existing solar arrays, formerly used as solar power plants or as thermal test facilities, like the National Solar Thermal Test Facility (NSTTF) near Albuquerque, New Mexico (106°.5 W, 35° N). The principle is the same as for the stand-alone telescopes described in previous paragraphs, but in this case the reflectors (heliostats) are larger, since they were designed to collect solar radiation (see Fig. 1.8). In addition, the camera is located in a tower which contains all the optics that are needed to focus the light — reflected from the mirrors — on the PMTs. The collection area of this experiment is enormous compared to a stand-alone telescope's. For example, the CELESTE experiment, located on the French Pyrenees at Thémis (2° W, 42.5° N), uses 40 mirrors with a total collection area of 2,000 m^2, which were formerly part of a solar power plant. Due to their large reflective areas, such arrays can collect a larger amount of photons than the smaller Cherenkov reflectors. Hence, they tend to be more sensitive in the low-energy range, and experiment thresholds of 50 GeV are not uncommon.

Background rejection with such experiments is achieved by reconstructing the shower direction using the individual arrival times of the Cherenkov photons at each heliostat: a process called wavefront sampling. Also, the shower energy is estimated using the total light deposited on the PMTs combined with information about the light intensity distribution around the shower axis at ground level. Discrimination between reconstructed γ-ray- and hadron-initiated showers is then achieved via selection cuts on the temporal width and total light (total charge on the PMTs) of the shower.

Although wavefront sampling can reject a large fraction of the background events, solar arrays lack the discrimination efficiency offered by the high-resolution cameras of Cherenkov imaging detectors. Nevertheless, at energies below 100 GeV, where imaging suffers from low light levels (see section 4.2.5), the gap between the techniques is reduced. Moreover, the large separation between the mirrors in solar arrays, as well as their narrow FoV, makes them insensitive to local muon showers that plague low-energy, single-telescope observations with imaging detectors; muon contamination is, however, not an issue with stereoscopic observations, as will be explained in chapter 4.

Solar array experiments are previously established constructions, which makes them a very cheap way to observe γ rays. Unfortunately, this implies that the mirrors

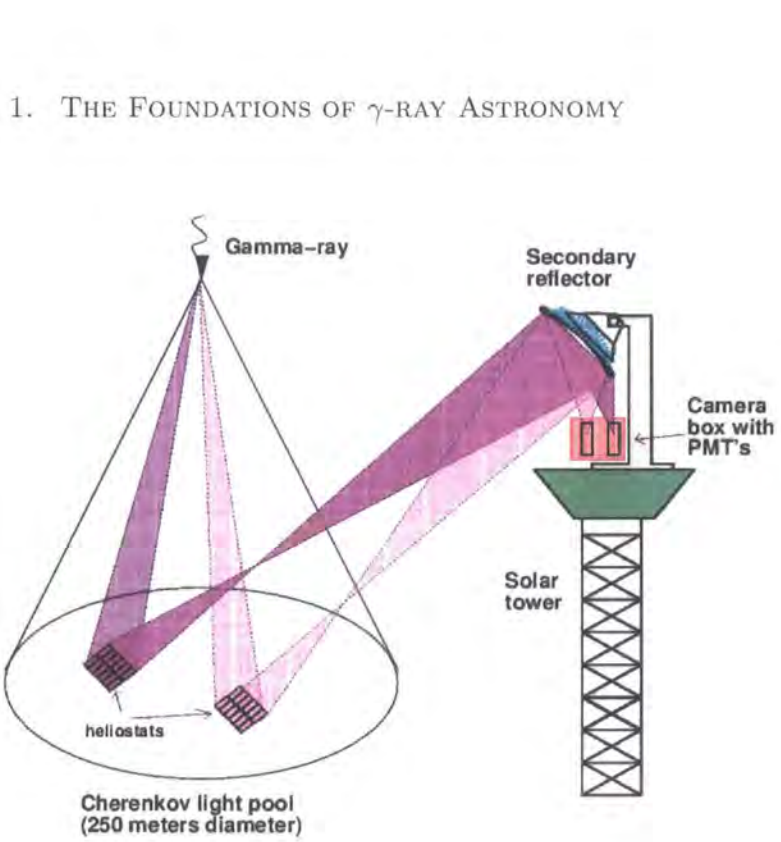


Figure 1.8: (from [54]) The layout of how an array of heliostat mirrors can be used for the detection of VHE γ rays.

are not optimised for the reflection of Cherenkov light, and therefore their performance in terms of reflectivity at short wavelengths is inferior to that of Cherenkov telescopes. To make matters worse, having the focus fixed on a tower means that the light is reflected at different angles depending on the location of the source with respect to the optical axes of the mirrors; this makes the arrays' light-collection efficiency prone to aberrations [52].

Since 2004, the STACEE collaboration has been using all 64 heliostats of the NSTTF facility, and its DAQ (Data Acquisition System) has been upgraded to an acquisition rate of 1 GSample s^{-1} channel^{-1}. As a result, its fast electronics will provide an even better parametrisation of the air showers, which will assist the γ/hadron discrimination and bring the experiment's threshold below 100 GeV [53].

1.4 Ultra High Energy Detectors

1.4.1 Detection Techniques

A small fraction of the γ rays that hit the Earth can be so energetic that they produce air showers whose secondary particles can survive the multiple scattering through the atmosphere and reach the ground. It is possible to detect these particles, provided there is a sufficiently large amount of them at ground level. Air showers that have relatively low energy ($\sim$ 1 TeV) are more abundant, but most of their secondary particles become thermal long before ground level. On the other hand, the very energetic showers ($\sim$ 1 PeV) are very scarce, but their products reach the ground at large numbers. More specifically, simulated showers from 1-TeV and 1-PeV γ rays result in 60 and 6×10^5 particles at 2.3 km altitude, respectively [1]. Hence, experiments with large collection areas ($\sim 10^4$ m^2) designed to capture these particles can be used for UHE γ-ray astronomy.

There exist several UHE γ-ray experiments, and they are mainly divided into two categories: those which directly detect the EAS particles and those which detect the Cherenkov light that the secondary particles produce when they enter a large water pond — which is part of the detector. The discrimination between γ rays and cosmic rays in those experiments is based on the anisotropies in the event numbers across the sky, and therefore good angular resolution is an asset. A clear advantage over the previously mentioned telescopes is the continuous monitoring of the entire overhead sky, as opposed to Cherenkov detectors which require moonless nights to operate.

The first category of UHE detectors, the particle air shower arrays, are usually equipped with scintillator counters located at high altitude. The reconstruction of the shower direction is performed by mapping the shower front using the arrival times at the different scintillators of the array. Unfortunately, the shower-front curvature is most of the time irregular, which introduces some uncertainty in the shower direction and consequently results in poor angular resolution ($\sim 1^\circ$). Nevertheless, the performance of these arrays becomes better with increasing shower size.

Water pond detectors, on the other hand, capture the secondary particles created in an air shower and record the Cherenkov light emitted as the particles pass through their large water mass. Placed on the pond's floor, there are several PMTs which are spaced apart at such distances as to record all Cherenkov photons created in the pond. The direction and energy of the shower can then be derived simply by means of timing information and amount of collected light on the PMTs, respectively.

1.4.2 Instruments

One example of a UHE detector is the Tibet scintillator array in China (90°.5 E, 30° N). It is built at an altitude of 4,300 m a.s.l. and has been upgraded many times in 13 years of operation. The latest version of the experiment, Tibet-III, consists of 533 scintillation counters with a 3-TeV energy threshold for proton-initiated showers. Considering the relatively poor angular resolution of the array ($\approx$ 0.9° above 3 TeV), the experiment has been quite successful in detecting γ rays from the Crab nebula and Mkn 421. In addition, the results from all Tibet arrays were cross-correlated with those from stand-alone VHE telescopes, like Whipple, which have better sensitivity and angular resolution ($\sim$ 2 arcmin at 3 TeV). This effort aimed at the more accurate localisation of the γ-ray sources seen with Tibet [55],[44],[56].

A completely different type of UHE detector is MILAGRO (Multiple Institution Los Alamos Gamma Ray Observatory) in Los Alamos, New Mexico (106°.7 W, 36° N). MILAGRO is the only active water pond experiment to date. Its pond can hold almost 40,000 m^3 of water and, unlike other detectors of its kind, has been built to detect all sorts of particles (e.g. $\mu^{\pm}$, $e^{\pm}$ and p^{+}). The detection of Cherenkov light from those particles is due to three PMT layers that comprise a total of 723 PMTs installed inside the water pond. Each PMT layer is responsible for the detection of a different particle species (Fig 1.9). Although MILAGRO has been designed to detect air showers up to PeV energies, its energy threshold ($\sim$ 250 GeV) is competitive with those of Air Cherenkov telescopes, like H.E.S.S., Whipple, etc. In 2002, the experiment was enhanced with complementary surrounding detectors that have lowered its energy threshold (50 m^2 effective area at $\sim$ 100 GeV) and increased its energy resolution (50%) and angular resolution (0.45°) [57]. In that way, the γ/hadron discrimination will be more efficient. This additional array of detectors will consist of 175 smaller water ponds, and together with the central pond it will cover an area of 40,000 m^2.

Lacking the imaging capabilities of VHE atmospheric detectors means that UHE experiments solely rely on the shower direction to discriminate between background and signal. Hence they are less sensitive than the former and their positive results have been very few and far in between.

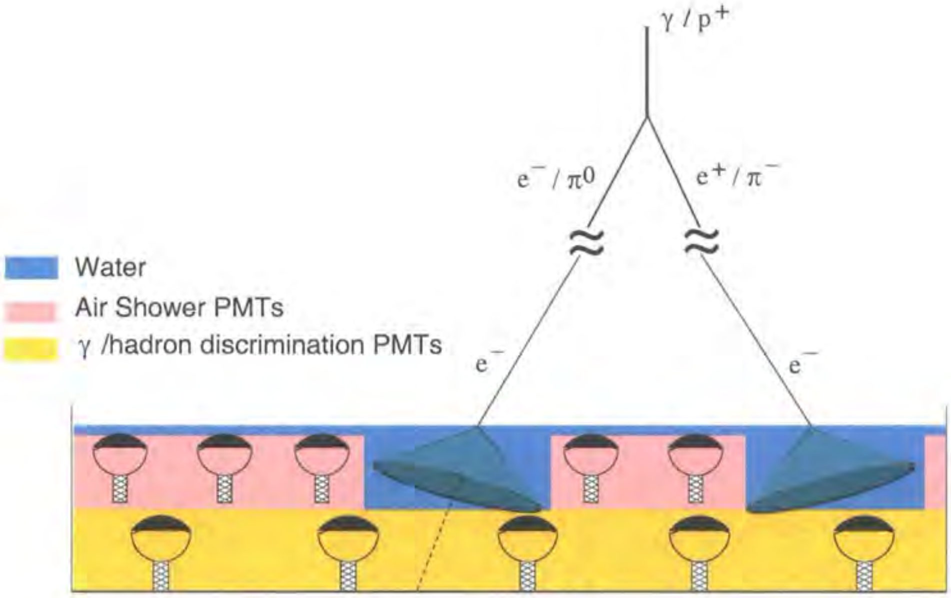


Figure 1.9: Basic principle of the operation of a water-pond γ-ray detector.

1.5 Summary

To summarise, one can say that each technique lacks certain features but possesses certain qualities compared to the rest. Space-borne experiments like EGRET are very expensive and cease to become useful in the VHE domain where the Air Cherenkov technique takes over. Air Cherenkov Telescopes have excellent angular resolution and can provide exciting results for suspected γ-ray sources in the GeV–TeV energy range, but their duty cycle is short — even under perfect skies, an ACT can only operate 20% of the time — which is not an issue for experiments like MILAGRO.

However, despite the partial overlap between the different experiments in terms of energy coverage (see Fig. 1.10), pointing ability, etc., γ-ray observations from HE up to UHE will complement each other and produce a complete picture of the processes of high-energy emission. Some of the contemporary γ-ray experiments together with their basic properties are listed in Table 1.4. Also, the locations of some of the experiments are shown in Fig. 1.11

Figure 1.10: (from [58]) Energy overlap between the various kinds of detection techniques.

Experiment	Energy domain	Type	Altitude (km)	Reference
INTEGRAL	20 keV – 10 MeV	Satellite	60,000	[59]
SWIFT (BAT)	15 keV – 150 keV	Satellite	600,000	[20]
VERITAS (Whipple)	100 GeV – 10 TeV	Air Cherenkov	2.3	[35]
CANGAROO-III	> 100 GeV	Air Cherenkov	0.165	[60]
H.E.S.S. (Phase I)	> 100 GeV	Air Cherenkov	1.8	[61]
MAGIC	> 30 GeV	Air Cherenkov	2.2	[48]
STACEE	50 – 500 GeV	Solar Array	1.7	[53]
MILAGRO	$\gtrsim 100$ GeV	Water Pond	2.64	[62]
GRAPES-II/III	100 TeV – 10 PeV	Counter Array	2.2	[63]
ARGO-YBJ	100 GeV – 20 TeV	Counter Array	4.3	[64]
Tibet-III	> 1 TeV	Counter Array	4.3	[65]

Table 1.4: Active γ-ray experiments as of the time of this thesis.

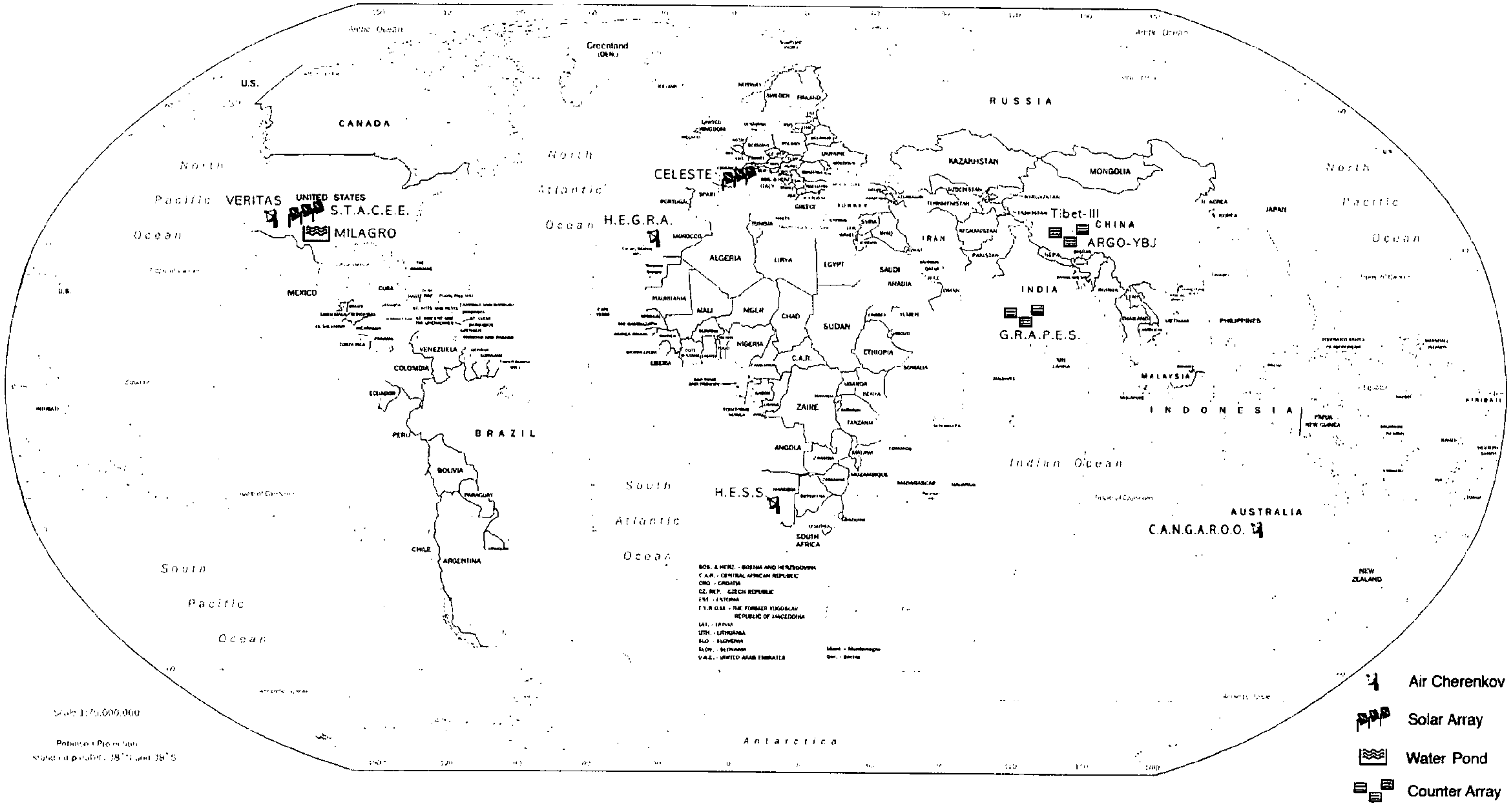


Figure 1.11: A Robinson projection showing the locations of some of the most well-known high-energy experiments.

1.6 A Review of Very High Energy γ-ray Sources and Recent Results

1.6.1 Introduction

In the period 1991–1995, the EGRET all-sky survey detected 271 γ-ray sources above 100 MeV [14]. Amongst them were 5 pulsars, one solar flare, probably the radio galaxy Cen A, and 93 blazars of which 66 were detected at a high confidence level. But also EGRET detected another 170 sources that could not be identified with any previously discovered objects, at any wavelength. Therefore, ground-based experiments had an abundance of potential VHE sources to study and to map their spectra, in the GeV–TeV range.

The years after *CGRO* saw the appearance of various international collaborations (e.g. VERITAS, CANGAROO, H.E.S.S. and MAGIC), whose ground-based experiments have extended the EGRET observations to the VHE range and, in a few cases, brought to light previously unknown high-energy sources. The following paragraphs provide a description of the main classes of VHE source and present the highlights from each collaboration's results in the period 2000–2005.

1.6.2 Galactic Sources

Supernova Remnants

The Galaxy offers plenty of astrophysical sources for VHE γ-ray astronomy, amongst which there are particle accelerators producing γ rays via various particle interaction mechanisms, and annihilators directly producing VHE γ rays. Amongst them are Supernova Remnants (SNRs), which are the remains of a catastrophic explosion at the end of a giant star's life. Based on the current observed sample, SNRs manifest themselves in three different ways:

> **Plerionic SNRs**
>
> Plerionic SNRs appear as blobs of expanding matter within which lies a pulsar. The pulsar's radiation pressure (i.e. the pulsar wind) pushes against the SNR's gas, which leads to the formation of a bubble around the pulsar that encloses a wind of relativistic particles: i.e. a pulsar wind nebula (PWN). In some cases, the pulsar wind leads to the formation of observable jets, which blow material out of the remnant. The base model for such SNRs, and a standard candle for

γ-ray astronomy, is the Crab nebula. VHE γ rays are generated within the nebula itself via the inverse Compton (IC) and pion decay mechanisms, and have been detected by various experiments (see e.g. [66]).

Another example of a plerionic SNR is the unusually shaped — as seen in radio waves — **MSH 15-52**, which was detected with H.E.S.S. in 2004, after 22 h of observations [67]. Previous observations of this PWN with CANGAROO-I had yielded only a marginal signal (4.1 σ) above 1.9 TeV, an estimated 10% of the Crab nebula's flux [68]. The extended, elliptical shape of MSH 15-52 above 900 GeV (semi-major axis $\sim$ 6 arcmin) was consistent with *ROSAT*'s X-ray map in the 0.6–2.1 keV range; and they both implied a pulsar-driven nebular emission. The energy source is identified as the 150-ms PSR B1509$-$58, which is located within the PWN. The detected TeV emission from 280 GeV up to 40 TeV ($\approx$ 15% of the Crab nebula's above the same threshold) could be fitted with a single power law ($\nu = 2.27$), and the emission mechanism was explained as inverse Compton interaction between pulsar-wind electrons and soft photons. This detection provided the first image of an extended PWN in TeV energies.

Amongst the possible candidates for a second detection of an extended PWN is **HESS J1825$-$137**, which was discovered with H.E.S.S. in the inner Galactic survey of 2004 [69]. There are currently a number of arguments for an association between the H.E.S.S. source and PWN G18.0$-$0.7 [70]. This nebula is driven by the pulsar PSR J1826$-$1334, which lies at the edge of a 10-arcmin radius around the centre of the H.E.S.S. source; the latter extends to $\approx 0.5^\circ$ in diameter. *XMM-Newton* and *ROSAT* observations, on the other hand, had detected a synchrotron component from G18.0$-$0.7, which extends to only 5 arcmin radius; and therefore HESS J1825$-$137 is a much larger source. Despite this inconsistency, the positional coincidence of HESS J1825$-$137 with G18.0$-$0.7 cannot be ignored; and the derived spectral index (ν = 2.2–2.6) can be well-explained by means of inverse Compton emission from electrons injected by the pulsar. Moreover, the size discrepancy can be lifted by arguing that the distance to the PWN (3.9 $\pm$ 0.4 kpc) could be erroneous; or, indeed, that the size seen in TeV energies is larger due to a diffuse electron component (past the size of the synchrotron nebula seen in X-rays), which can still emit via inverse Compton but little does so via synchrotron.

Other examples of plerions are the Vela SNR, which hosts the Vela pulsar, and the SNR G343.1−2.3, which is believed to host the radio pulsar PSR B1706−44 [71],[34],[72].

Shell-type SNRs

Alternatively, SNRs can appear as expanding shells of gas, without having a visible pulsar within their volume. In this case, they are characterised as shell-type SNRs. They appear as having a ring structure because of the limb brightening that is caused by the denser column of hot gas along the line-of-sight: the latter being tangential to the shell's boundaries. The gas temperature rises as the initial shock wave from the explosion compresses the stationary interstellar matter, during its radial propagation through space. It is believed that the TeV emission in shell-type SNRs is generated via inverse Compton and non-thermal Bremsstrahlung processes from energetic electrons and possibly via π^0 decay from proton-proton collisions in the densest areas of the nebula [73]. A classic example of a shell-type SNR is **Cassiopeia A** (Cas A), which is a northern-sky SNR and the strongest observed radio source. Nevertheless, at TeV energies Cas A appears as a very weak emitter, with only 3.3% of the Crab nebula's flux above 1 TeV [74].

The first confirmation of a TeV shell-type SNR that exhibits clear signs of particle acceleration within its shell came from observations of **RX J1713.7−3946** in 2003 with the H.E.S.S. Previous observations of this object, which had already been mapped in the 1–3 keV range with the *ASCA* X-ray satellite, clearly show an extended region of non-thermal, X-ray emission. Part of the SNR had also been investigated in 1998 with the original 3.8-m CANGAROO-I telescope [75]: most of the emission appeared to be coming from the NW rim of the SNR, which was detected at the 5.6-σ level above 1 TeV. A follow-up paper, published in *Nature*, reported an energy spectrum from that part of the source with spectral index 2.84; the spectrum implied a pion decay mechanism of TeV γ rays [76].

Unlike the CANGAROO detection of only a part of the SNR, H.E.S.S. mapped the whole extent of RX J1713.7−3946 with a resolution of 3 arcmin. The emission sites within the object boundaries correlated well with their X-ray counterparts that were observed with *ASCA*. In the 100 GeV–10 TeV range, the energy spectrum from the entire source was well represented by a power law with index 2.19, which was in disagreement with the CANGAROO result from the NW rim. But, most importantly, this discovery provided unequivocal evidence of particle acceleration (up to particle energies of 100 TeV) in the expanding shells of SNRs [73].

RX J0852.0−4622 is another shell-type SNR, very similar to RX J1713.7−3946, which was detected with CANGAROO-II, in data from various periods during 2001, 2002 and 2003. The observations were focused on the NW rim of the remnant, corresponding to the peak of X-ray emission, and resulted in a significant detection at the level of 6 σ, above 500 GeV [77]. At 1 TeV, the estimated flux from the γ-ray source was 12% of that from the Crab nebula. Finally, the derived differential flux spectrum from the investigated region was $dN/dE \propto E^{-4.3}$, which is quite steep compared to the Crab nebula's $dN/dE \propto E^{-2.6}$. However, this was attributed to the large uncertainty in the energy estimation.

H.E.S.S. also detected RX J0852.0−4622, at the 12-σ significance level, in only 3.2 h of data. The source appeared as extended, with $\sim 1^\circ$ radius [78] (see Fig. 1.12). A more confined observation to the centre of the source did not reveal a central object, which had otherwise been seen in deep X-ray observations with *ASCA*, *BeppoSAX* and *Chandra* [79],[80],[81]. Nevertheless, the mapped SNR shell correlated with the *ROSAT* flux contours, although lengthier observations are required in order to have a more detailed γ-ray morphology. As is the case with RX J1713.7−3946, the measured spectrum of RX J0852.0−4622 ($\nu = 2.1$, in 0.5–15 TeV) leads to π^0 decay as a more favourable scenario over inverse Compton; and hence it provides further evidence for cosmic-ray acceleration at the expanding shells of SNRs. However it is agreed that a better estimation — via future observations — of the magnetic field value, which is the critical factor, is required to have a more definite conclusion.

In early 2004, with only two of the additional three telescopes of the CANGAROO-III phase ready, the collaboration investigated RX J0852.0−4622 anew. This time the detection was at the 7-σ level, and the preliminary spectrum was slightly softer than the one of H.E.S.S. — although more consistent with the latter than that of CANGAROO-II [82].

Early observations (2003–2004) of the shell-type SNR **SN 1006** with the H.E.S.S. array failed to detect γ-ray emission above 110 GeV [83]. However, the non-detection with a sensitive instrument such as H.E.S.S. had major implications on previous results from this object. Observations of this SNR with H.E.G.R.A. in 1999–2001 resulted in a $\approx$ 5-σ detection with a flux estimate of $F(> 18 \text{ TeV}) = 2.5 \pm 0.5 \times 10^{-13}$ cm^{-2} s^{-1} [84]. Also, the CANGAROO-I 3.8-m telescope had observed SN 1006 in 1996 and 1997, and the collaboration reported a signal for both those periods: the 1996 data yielded a flux equal to $F(> 3 \pm 0.9 \text{ TeV}) = 2.4 \pm 0.44 \times 10^{-12}$ cm^{-2} s^{-1}, and the 1997 data, $F(> 1.7 \pm 0.5 \text{ TeV}) = 4.6 \pm 0.6 \times 10^{-12}$ cm^{-2} s^{-1} [85]. Later, in 2000, the same collaboration used the 10-m CANGAROO-II telescope to re-investigate the emission from SN 1006. The source was confidently detected again (at 6.5 σ) and the reported flux was $F(> 1 \text{ TeV}) = 6 \pm 1.0 \times 10^{-12}$ cm^{-2} s^{-1} [86]. Based on 24.5 h of data on this source, H.E.S.S. derived upper limits that were an order of magnitude lower than the CANGAROO 1996–1997 fluxes. More specifically, the calculated upper limits on the flux from the NE rim of SN 1006, which matched CANGAROO's investigated location and were above CANGAROO's energy threshold, were in the range $F_{\text{ul}}(> 1.7 \text{ TeV}) = 0.34\text{–}0.49 \times 10^{-12}$ cm^{-2} s^{-1}. Furthermore, a direct comparison with the H.E.G.R.A. flux required a high-threshold cut (> 18 TeV), which resulted however in zero events. Considering the Poissonian error on this measurement (a mere 7 counts), H.E.G.R.A.'s result appeared consistent with the signal absence with H.E.S.S, within the quoted $\geq$ 35% systematic error. Finally, the CANGAROO-II results, which were reported as consistent with the 1996–1997 observations, were also considered inconsistent with the H.E.S.S. upper limits. However, follow-up observations of the same regions of SN 1006 with the more sensitive CANGAROO-III verified the absence of signal [82].

Following the H.E.S.S. collaboration's discovery of **HESS J1813$-$178** in the inner Galactic survey of 2004, the MAGIC telescope was used to observe this high-interest source — which has been associated with the shell-type SNR G12.82$-$0.02, at $\approx$ 94% chance probability [87]. The observations confirmed this source as a VHE emitter (10.6 σ in 25 h), and furthermore the spectral analysis revealed a hard spectrum ($\nu \approx 2.1$), which was consistent with that of H.E.S.S. The spectral distribution at VHE, combined with radio, X-ray and lower-energy γ rays strengthen even more the association with the SNR: the particle acceleration in those regions is expected to lead to hard VHE tails (see [88] and references therein).

Composite SNRs

Finally, the morphology of SNRs can be a mixture of both the above types. They may also differ depending on the wavelength at which they are observed. This type of SNR is referred to as composite and can display, for example, a Crab-like morphology at high energies, whereas appear as shell-type at radio wavelengths.

A well-known composite SNR is **G0.9+0.1**, which is situated near the Galactic centre ($\sim 1^\circ$ away from Sgr A*). The radio map of this source shows an admixture of a compact, central core and an expanding, shell-type structure [89]. The VHE emission from G0.9+0.1 is thought to originate from the interaction of accelerated leptons with ambient, low-energy (IR, microwave) photons via inverse Compton. In this scenario, the leptons are accelerated at the termination shock between the relativistic particle wind — instigated by a central compact object: e.g. a pulsar — and the surrounding medium. Observations with H.E.S.S. supported the above image, as the source appeared point-like in the field of view, which suggests emission from a central compact object.

Amongst the experiments that have observed G0.9+0.1 at TeV energies has also been H.E.G.R.A., which only managed to constrain the flux below $F_{\mathrm{ul}}(> 4.2\ \mathrm{TeV}) = 4.6 \times 10^{-12}\ \mathrm{cm}^{-2}\ \mathrm{s}^{-1}$ [90]. On the other hand, H.E.S.S. achieved a clear detection at the level of 13 σ, after having observed this source for 60 h [91]. The estimated flux above 200 GeV was $F(> 200\ \mathrm{GeV}) = 5.7 \pm 0.7 \pm 1.2 \times 10^{-12}\ \mathrm{cm}^{-2}\ \mathrm{s}^{-1}$ ($\approx 2\%$ of the Crab nebula's flux), which makes this source one of the weakest in the TeV sky. Based on the 8.5 kpc distance to this object, the luminosity in VHE γ rays was $\approx 50\%$ that from the Crab nebula: i.e. $L_{\mathrm{G0.9+0.1}} \sim 2 \times 10^{34}\ \mathrm{erg\ s}^{-1}$.

Table 1.5 presents an overview of the measured properties from recent VHE observations of SNRs.

Source	Type	T (h)	σ	$E_{\rm th}$ (TeV)	$F(>E_{\rm th})$ (cm^{-2} s^{-1})	ν	Reference
MSH 15-52	PWN	22	25	0.28	$\approx 2.25 \times 10^{-11}$	$2.27 \pm 0.03 \pm 0.2$	[67]
PSR B1706−44 (nebula)	PWN?	14.3	2	0.35	$< 1.4 \times 10^{-12}$	*2.5*	[92]
Cas A	shell-type	232	5	1	$5.8 \pm 1.2 \pm 1.2 \times 10^{-13}$	$2.5 \pm 0.4 \pm 0.1$	[74]
SN 1006	shell-type	24.5	3	1.7	< 0.34–0.49×10^{-12}	*2.0*	[83]
RX J1713.7−3946	shell-type	18.1	20	1	$1.4 \pm 0.1 \pm 0.3 \times 10^{-11}$	$2.19 \pm 0.09 \pm 0.1$	[73]
RX J0852.0−4622	shell-type	3.2	12	1	$1.9 \pm 0.3 \pm 0.6 \times 10^{-11}$	$2.1 \pm 0.1 \pm 0.2$	[78]
G0.9+0.1	composite	50	13	0.2	$5.7 \pm 0.7 \pm 1.2 \times 10^{-12}$	$2.4 \pm 0.11 \pm 0.2$	[91]

Table 1.5: Summary of recent results from observations of SNRs. Apart from Cas A, which was detected with H.E.G.R.A., the rest of the results correspond to H.E.S.S. observations. The third column presents the exposure times (T) of the data sets that were used in the respective analyses (livetime); and the fourth, the resulting statistical significance of the signal in standard deviations above the background. Where only an upper limit is reported (indicated with the inequality sign under column 6), the significance shown corresponds to the confidence level of that upper limit. The sixth column shows the integral fluxes, $F(>E_{\rm th})$, above the respective energy thresholds, $E_{\rm th}$ (column 5). Finally, the seventh column shows the spectral index of the differential power-law spectrum as was fitted across each investigated range. Where one or more errors are stated, the first corresponds to the statistical uncertainty and the second, to the systematic one. Also, the flux upper limits — when no-detection has been reported — are indicated with the inequality sign; the corresponding spectral index was assumed prior to the upper limit calculation.

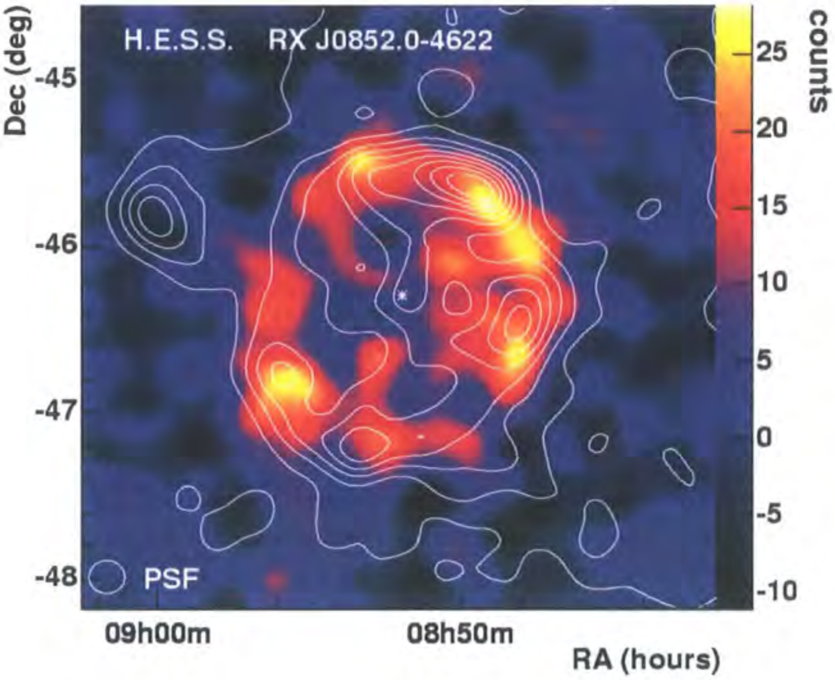


Figure 1.12: (from [78]) γ-ray map of the shell-type SNR RX J0852.0−4622, as was observed with the H.E.S.S. telescopes. The x and y axes show the J2000 Right Ascension (R.A.) and Declination (Dec) co-ordinates, respectively. The white circle in the lower left corner indicates the angular resolution of the instrument (0.1°); map features smaller than that are not real. Overlaid on this map are X-ray contours from observations of this object with the *ROSAT* X-ray satellite; only X-ray photons above 1.3 KeV have been considered for its construction. The vertical, graduated colour band shows the colours corresponding to the number of excess γ rays after the subtraction of the background. The star at the centre of the map shows the position of the central compact object detected with various X-ray experiments (see text).

Pulsars

Since pulsars are the main topic of this thesis, we will only mention a few characteristic examples here. More detail can be found in subsequent chapters.

Pulsars can be categorised according to age, energy range, period, etc. Although there are representative samples from each category, the distinction is not strict, and many of them belong in more than one categories. So far there has not been an established detection of pulsed emission in the VHE regime, and all the available information on their high-energy properties comes from EGRET observations up to a few GeV. At TeV energies, there have been reports of pulsed emission from the Crab pulsar, but more recent experiments have not confirmed the previous results [93],[94],[95].

MAGIC is one example of VHE experiment which has been active in pulsar observations; particularly, it has been used to investigate millisecond pulsars. In October 2004, the MAGIC collaboration collected 6 h of data from **PSR B1957+20**, one of the fastest-spinning pulsars known, and $\sim$ 13 h from **PSR J0218+4232**, a luminous 2.3-ms radio pulsar which is positionally coincident with the EGRET unidentified source 3EG J0222+4253. Both pulsars are part of a binary configuration and their systems have been detected in X-rays (see e.g. [96]) — which are thought to arise from the interactions between the pulsar wind and the stellar companion's photon field. Moreover, these systems are also expected to emit γ rays generated not only from the pulsar–star interaction but also directly from the millisecond pulsars' magnetospheres. In fact, PSR J0218+4232 may already be regarded as a HE γ-ray emitter — if its association with 3EG J0222+4253 is true — since pulsed emission from the EGRET source, up to 1 GeV, has already been detected with *CGRO*, at the 3.5-σ level [97]. Nevertheless, above 115 GeV, the data collected with MAGIC were subjected to a number of sensitive periodicity tests, able to resolve even the most complex pulse profiles, but they were also tested for a continuous VHE emission. In none of the cases was there a significant excess, and upper limits were derived: with regards to the continuous emission, the upper limits to a 95% confidence level for PSR B1957+20 and PSR J0218+4232 were $F_{\rm ul}(> 115~{\rm GeV}) = 4.5 \times 10^{-12}$ and $1.3 \times 10^{-12}~{\rm cm}^{-2}~{\rm s}^{-1}$, respectively. An extended review of pulsar observations with H.E.S.S. can be found in section 6.2.

Binary Systems

Another target for high-energy observations are accreting binaries. Typical binary systems that are expected to emit VHE γ rays include X-ray binaries, X-ray transients, microquasars, etc. X-ray binaries involve a compact object (i.e. neutron star/white dwarf) or a black hole that accretes matter from a main sequence or post–main sequence star. Depending on the mass of the stellar companion, X-ray binaries are classified as Low-Mass X-Ray Binaries (LMXRBs), if the stellar companion has $M \sim M_{\odot}$, or High-Mass X-Ray Binaries (HMXRBs), if the stellar companion is a giant O-type or B-type star. As their name suggests, these objects are predominantly emitting X-rays. The emission is produced either through the stellar wind interaction with the compact object, which is typically the case for HMXRBs, or via an accretion mechanism, which is often seen in LMXRBs [98].

An example of an accreting HMXRB involving a pulsar and an O-type supergiant is **Cen X-3**. This system is known to possess a complex temporal variability: the pulsar that accretes matter from the stellar companion has a rotation period of 4.8 s and an orbital, of 2.1 d [99],[100]. These periodicities affect the X-ray emission, which shows a synchronous modulation. Nevertheless, in the VHE range, there have only been a few indications of modulated emission; these have not been confirmed by modern experiments.

Early observations of Cen X-3 by the Durham and Potchefstroom groups — with non-imaging telescopes — detected sporadic outbursts of TeV emission at the pulsar's frequency [101],[102]. The former group, using the Mark 3 detector (see section 1.3.4), detected a ≈ 4.5-σ signal during a small fraction (5%) of the pulsar's orbit corresponding to the ascending node. An independent confirmation of the VHE emission came from observations with the Potchefstroom telescope during 1986–1989: they reported a 4.8-s periodic signal, also confined to the same orbital phase range, but at the lower significance of 3 σ. Later investigation of this object with the more sensitive Durham Mark 6 imaging telescope, at energies above 400 GeV, resulted in a 4.5-σ detection, which was however constant: i.e. it was not modulated with either the orbital or the pulsar's period [103].

Another good example of a cataclysmic X-ray binary is **Cyg X-3**. It is believed to be a system comprising a pulsar that revolves around a Wolf–Rayet star [104],[105]. Although earlier experiments had reported a significant excess from the direction of Cyg X-3, more recent efforts have failed to reproduce those results: see [106],[107],[108],[109] and [110], but also [111],[112] and [113].

A particular class of XRBs exhibits powerful jet emission that resembles, on a smaller scale, that from Active Galactic Nuclei (AGN) — also known as quasars. A description of AGN is given in the following sections, when we discuss the extragalactic VHE sources; but, in brief, they are the cores of luminous, high-redshift galaxies from which jets of ultra-relativistic particles (along with intense electromagnetic emission) emanate; the energy source of AGN is thought to be an accreting supermassive black hole. Observations have shown that a number of XRBs are sources of powerful beamed emission that seems to arise from a central compact object or a black hole that accretes material from a companion star. In those systems, the accretion mechanism is so violent that it results in the formation of ultra-relativistic jets of particles. There is supporting evidence for electron acceleration processes somewhere along the jet, which could lead to VHE γ-ray emission via the inverse Compton process, given a seed photon field [114]. Due to the high Lorentz factors of the accelerated particles in the jets — which can result in particle velocities of $0.98c$ — the generated emission is Doppler-boosted to potentially TeV energies. The similarity between the features observed in quasars and those systems, albeit displayed on a much smaller scale in the latter systems, has given them the name microquasars.

However, VHE emission from the direction of known microquasars has only been recently associated with high confidence, with these objects: the microquasar **LS 5039** was associated, at a 3-σ confidence level, with the H.E.S.S. source HESS J1826−148; and the measurements revealed persistent VHE emission above 250 GeV, extending to 4 TeV. The significance of the detection was 8.2 σ, and the integral flux, $F(> 250 \text{ GeV}) = 5.1 \pm 0.8 \pm 1.3 \times 10^{-12}$ cm^{-2} s^{-1} [115].

Amongst the known binary-system configurations, there are a few that are unique. One example is the binary radio pulsar **PSR B1259−63** that orbits the Be star SS 2883, which is the only such configuration known so far (but see also section 6.2). Its highly eccentric orbit together with the Be star's extended disc of emission conspires towards a wind interaction between the pulsar and its companion, during periastron. This interaction produces VHE γ rays that are detectable in the GeV–TeV energies [116]: observations of the system in February–June 2004, during the pulsar's pre- and post-periastron phases, showed a significant excess of continuous γ-ray emission above 380 GeV, at a level higher than 13 σ [117]. Models describing the high-energy emission from this system around periastron had existed before H.E.S.S. However, despite the fact that this system had been observed in

X-rays and soft γ rays, the number of follow-up VHE observations was limited by the length of the system's orbit, which allows periastron observations only every $3\frac{1}{2}$ years [118],[119].

Dark Matter

The totally uncharted area of the Galactic dark matter could also be explored in the VHE range. The candidate location for an indirect detection of dark matter in our galaxy is the Milky Way halo, where dark matter density could be high enough to trigger neutralino annihilations at a detectable rate. However, the information currently available from observations, e.g. the total Galactic mass out to 100 kpc, can only constrain the dark matter density in the halo within 3 orders of magnitude [120]. Depending on the model parameters, the predicted γ-ray flux above 50 GeV and over a solid angle of 2×10^{-5} sr varies from 0.3×10^{-14} cm^{-2} s^{-1} to 108×10^{-14} cm^{-2} s^{-1} [121]. In addition, it is thought that the existence of a massive black hole in the centre of our galaxy would enhance the γ-ray emission as well as give rise to other signatures that are born from the annihilation process [122],[123]. These have the form of neutrino production as well as synchrotron radiation from the $e^{-}-e^{+}$ pairs produced in the annihilations.

The currently operating Cherenkov experiments, like MAGIC and H.E.S.S., are sensitive above 50–100 GeV: i.e. in the energy range where γ rays from dark matter annihilations are expected — especially from neutralinos with $\sim$ TeV masses. γ-ray emission from the Galactic centre has already been detected with most VHE detectors (see next section). However, the derived spectra extend to 20 TeV, while most dark-matter-annihilation models for a central compact source predict cut-offs below 10 TeV; this renders the validity of those models unlikely. Nevertheless, dark matter annihilation from an extended source in the direction of the Galactic centre (e.g. dark matter halos) cannot be excluded: in such case, the model predictions lead to spectral peaks in the 10–100 GeV range, which are consistent with the observed fluxes from EGRET [124],[125],[126].

Galactic Centre

The centre of our Galaxy ($l = 0, b = 0$) is a powerful source of γ rays. Assuming that the source of emission lies at 8.5 kpc distance, its γ-ray luminosity above 100 MeV ($\approx 2.2 \times 10^{37}$ erg s^{-1}) is roughly ten times higher than that of the Crab nebula [124].

But, until recently, observations in HE with EGRET could not resolve the emission as coming from a specific object, and many possibilities were proposed: (a) a peak in the Galactic diffuse γ-ray emission; (b) a massive black hole; (c) a group of radio-quiet pulsars situated in the foreground; (d) a supernova remnant or (e) dark matter annihilations [1]. Observations from the ground have resolved a source of VHE emission coincident with **Sgr A***: a massive black hole ($M_{\rm BH} \sim 3 \times 10^6\ M_\odot$ [127]) that is believed to trigger the production of VHE γ rays via particle acceleration in the accreted — by the black hole — stellar winds. The proposed mechanisms include the acceleration of electrons or protons in those winds or, alternatively, in nearby expanding supernova shells: e.g. Sgr A East. As was discussed in the previous section, dark matter annihilation in regions of dense neutralino production is also thought to be a possibility. Nevertheless, recent results from H.E.S.S. and MAGIC revealed VHE emission of *steady* nature. If the source of the emission is due to accretion-induced relativistic jets emanating from the black hole, then short-term fluctuations are expected. This scenario is similar to AGN emission, which clearly shows variability (see section 1.6.3). On the other hand, shock-induced acceleration of particles leads to more compatible behaviour: observations of the Crab nebula and G0.9+0.1 have shown steady emission over their investigated periods.

In the extended period 1995–2003, the Whipple telescope was used to observe the Galactic centre. 26 h of data from that source resulted in a 3.7-σ significance above 2.8 TeV, due to a source in the centre of the FoV. The integral flux above the energy threshold was $F(> 2.8\ \mathrm{TeV}) = 1.6 \pm 0.5 \pm 0.3 \times 10^{-12}\ \mathrm{cm}^{-2}\ \mathrm{s}^{-1}$ [128]. It should be noted that the derived flux above 2.8 TeV was a factor ~ 2 higher than that derived from subsequent observations with H.E.S.S. and MAGIC (see below).

In 2001–2002, the CANGAROO-II telescope was also used to observe the Galactic centre. In this case the conditions were more favourable, as the source could be observed at smaller Z.A.s. Data analysis showed a statistically significant excess above 250 GeV, and the observed profile was consistent with a point-like source [129].

As in the case of Whipple and CANGAROO, the interest in the Galactic centre was also high for H.E.S.S. In the summer of 2003, H.E.S.S. collected 11.8 h of data from **Sgr A***, which yielded a 9.2-σ detection above 165 GeV. The integral flux above the threshold was $F(> 165\ \mathrm{GeV}) = 1.82 \pm 0.22 \times 10^{-11}\ \mathrm{cm}^{-2}\ \mathrm{s}^{-1}$, whereas the flux above the Whipple telescope's threshold was $F(> 2.8\ \mathrm{TeV}) = 5.86 \pm 1.46 \times 10^{-13}\ \mathrm{cm}^{-2}\ \mathrm{s}^{-1}$. It is worth noting that the resulting energy spectrum was largely harder than that of CANGAROO's [126].

Very recently, in the period May–July 2005, it was MAGIC's turn to target the Galactic centre. Despite the less favourable location of MAGIC with respect to this source compared to H.E.S.S., a significant detection at the level of 7.3 σ after a 24-h exposure time allowed the construction of the source spectrum; the derived spectral index ($\nu \approx 2.2$) and differential flux ($2.9 \pm 0.6 \times 10^{-12}$ cm^{-2} s^{-1} TeV^{-1} at 1 TeV) were in good agreement with those measured with H.E.S.S. [124]. The emission extended up to ≈ 20 TeV.

Unidentified Sources

From Surveys

At the turn of the 21$^{\text{st}}$ century, the number of known VHE sources in the Galaxy were no more than 7 [130]. In 5 years this number has increased threefold [131]. Most of these sources were detected at the locations of previously known sources of lower-energy (radio, X-rays, soft γ rays) emission. However, more than a third of the Galactic VHE sample is the product of surveys.

One of the most important contributions of H.E.S.S. to our knowledge of the VHE sky was the recently performed survey of the inner Galactic plane. The survey resulted in 230 h of good quality data, collected in 2004, covering $l = \pm 30°$ in Galactic longitude and $b \pm 3°$ in latitude. The sensitivity of the survey above 200 GeV reached, on average, 2% of the Crab nebula's flux; and for weak sources above 250 GeV, H.E.S.S. was able to resolve emission regions that were less than 0.1° apart — at a 68% confidence level.

The H.E.S.S. inner Galactic survey resulted in the discovery of 14, previously unknown VHE sources, which were detected at a > 4-σ level; 8 of them were above 6 σ (see Fig. 1.13). A few of them have been confidently associated with SNRs previously seen in radio and X-rays, but for the majority there are less firm associations, or in some cases none at all. The present status of these associations is presented in Table 1.6, together with a few of the measured properties [69]. Here, we will briefly mention 5 of the sources whose association is unclear or nonexistent:

Amongst the unidentified sources was **HESS J1614−518**, which exhibits elliptical emission ($14' \times 9'$) and is one of the brightest sources discovered in the survey, with a flux above 200 GeV equal to 25% that from the Crab nabula. *Chandra* has searched for X-ray emission from the direction of this source but without success.

A weaker source (7% of the Crab nebula's flux), **HESS J1702−420**, was also not associated with any of the known objects in the field of view. The closest candidate is PSR J1702−4128, which could account for the observed VHE emission above 200 GeV, but that would also require an asymmetric PWN around the pulsar — which nevertheless has not been detected.

One of the new sources that had the benefit of 37.2 h exposure time — due to the nearby SNR RX J1713.7−3946 — was **HESS J1708−410**. However, no plausible counterpart has been found for this source whose flux above 200 GeV was equal to 4% that from the Crab nebula.

Finally, another two sources that have come out of the H.E.S.S. survey, **HESS J1745−303** and **HESS J1837−069**, were in positional coincidence with the unidentified EGRET source 3EG J1744−3011 and the *ASCA* X-ray source AX J1838.0−0655, respectively. However, if the former association is true, it requires a complex spectrum for this source, since a simple power-law extrapolation from sub-TeV energies (H.E.S.S.) to a few GeV (EGRET) leads to an integral flux (above 100 MeV) that is an order of magnitude lower than that detected with EGRET. On the other hand, HESS J1837−069 coincides with the brightest X-ray feature in *ASCA*'s map of the X-ray complex G25.5+0.0: AX J1838.0−0655. The nature of this source is unclear, but it is thought that it may be a synchrotron-emitting SNR like SN 1006, or a PWN.

The large FoV of the **MILAGRO** detector provides a full coverage of the overhead sky viewed from 36° latitude. In the period 2000–2003, the detector collected data from a narrow band across the Galactic plane ($|b| \leq 5^\circ$) [132]. The region with $40^\circ \leq l \leq 100^\circ$ showed an excess of γ rays from the background, at the 4.5-σ level. Although the signal strength translates to a marginal detection, it is important to note that the calculated flux above 3.5 TeV ($= 6.4 \pm 1.4 \pm 2.1 \times 10^{-11}$ cm^{-2} s^{-1} sr^{-1}) was consistent with an extrapolation of the HE observations of the same region with EGRET in the 10–30 GeV range. Consequently, the results indicated that a simple power-law

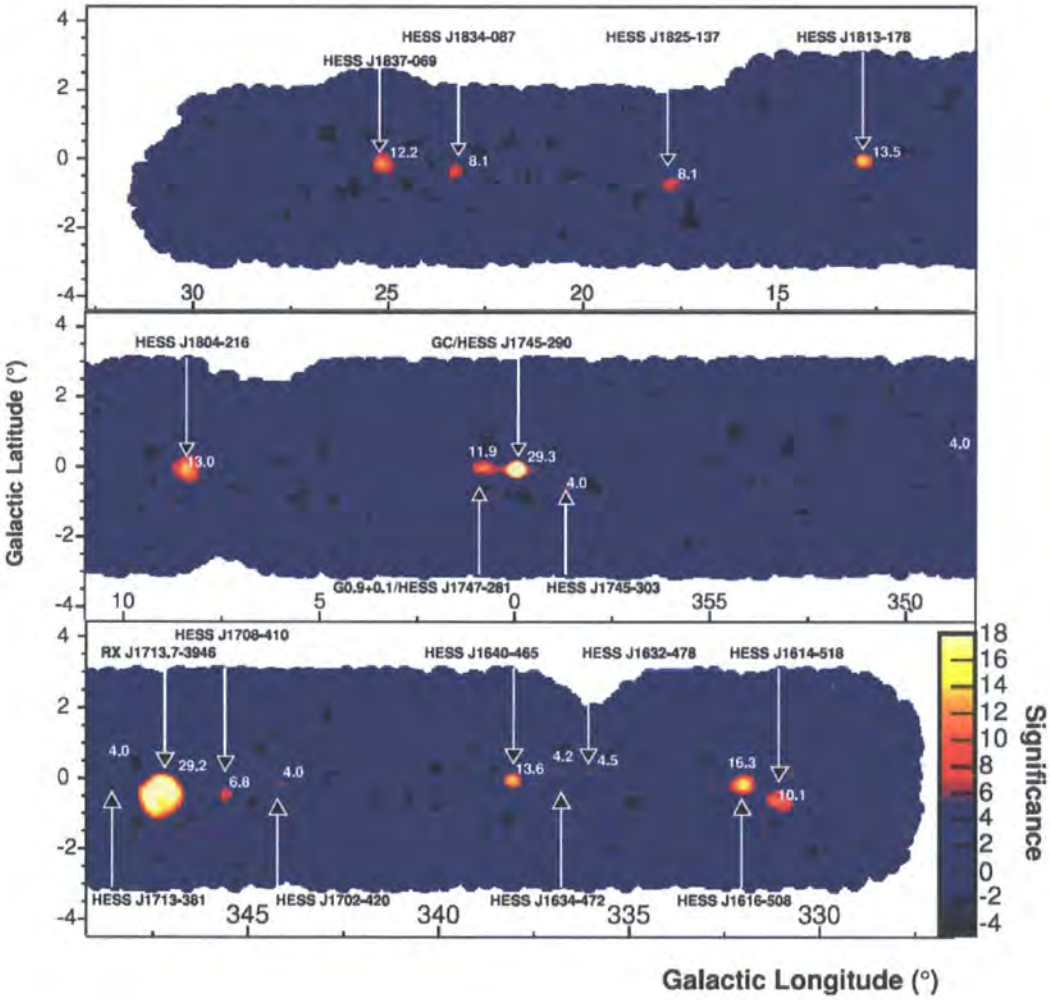


Figure 1.13: (from [69]) Significance map of the sources detected in the H.E.S.S. Galactic survey of 2004. The map also shows the shell-type SNR RX J1713.7−3946 and the Galactic centre region. The numbers next to each source give the significance of the corresponding detection in standard deviations above the background. The energy threshold of the observations was 250 GeV.

spectrum from EGRET's top energy range (where $\nu = 2.51$) to TeV energies (where $\nu = 2.61$) is feasible — where the softer spectrum in MILAGRO's range is explained by the steepening that is expected if π^0 decay is behind the TeV γ-ray production. The γ-ray signal from the mid-longitude Galactic region could not be resolved into a particular class of VHE source. Instead, the authors left the possibility open for the source to be any of the diffuse γ-ray emission, a number of unresolved point sources, or an extended γ-ray source. However, they remarked that the bins with the maximum excess are contained in the **Cygnus region** ($75° \leq l \leq 85°$): a well-known location for its pronounced emission in all wavelengths.

By Serendipity

Nevertheless, there are a few other unidentified sources which have been serendipitously discovered in the field of view of main targets.

On of these discoveries occurred during the detection of PSR B1259−63, which brought to light **HESS J1303−631**, in the camera's 5° FoV. Initial efforts to identify this source with a previously known object did not meet with success, and hence the source was given a name after the standard IAU nomenclature of the discovering experiment. Following this accidental discovery, the collaboration focused their observations on HESS J1303−631, and during February and June 2004 $\approx$ 55 h of good quality data were collected. The analysis resulted in a clear γ-ray signal (21-σ significance) from a rather extended source (0.16°). The flux was constant over the whole observation period and was calculated to be $(17 \pm 3)\%$ that from the Crab nebula above 380 GeV [133].

It is almost certain that future experiments like *GLAST* and *SWIFT* will try to re-investigate HESS J1303−631. However, the fact remains that it was not only the first time that this source was being detected in TeV γ rays, but also the first time that two γ-ray sources were observed simultaneously in the same FoV, which displays the value of H.E.S.S.'s wide FoV.

More recently, in 2005, the VERITAS collaboration confirmed the existence of a new TeV source, 0.6° north of Cyg X-3, after re-analysis of 1989–1990 Whipple archival data. This source, namely **TeV J2032+4130**, had already been seen from the Crimean Astrophysical Observatory and was also detected serendipitously with H.E.G.R.A. The latter had estimated the source flux to 3% of the Crab nebula's emission, whereas the new detection with the Whipple telescope placed it somewhat higher, at 12% of the Crab's [134],[135]. In both cases, the same source position was used, although the observations were conducted above different energy thresholds: those with H.E.G.R.A. were above 1 TeV, whereas the Whipple telescope's observations, above 0.6 TeV.

1.6.3 Extragalactic Sources

Starburst Galaxies

The high rate of star formation and an equally frequent supernova occurrence in starburst galaxies has given them their above name. These galaxies exhibit well-localised regions within their volume, in which supernova explosion rates are greatly enhanced ($\sim 10\times$) compared to those of normal galaxies [1]. Those regions, called "starburst regions", contain dense hot gas, and the photon densities are also much higher than those of the stellar nurseries in normal galaxies. Given the enhanced supernova rate in starburst galaxies, one can expect an abundance of accelerated cosmic rays produced in the explosions. A value of $E_c \geq 100$ eV cm^{-3} (~ 100 times higher than for the Galaxy) for the energy density of the produced cosmic rays is considered plausible in those regions [136]. Therefore, substantial VHE emission in the form of γ rays from π^0 decay should be expected. In addition, other processes, like inverse Compton upscattering of low-energy photons by relativistic electrons, or Bremsstrahlung emission from collisions of the latter with interstellar nuclei, are expected to play an important role in the production of VHE radiation.

Two typical examples of starburst galaxy are **NGC 253**, in the southern hemisphere, and **M 82**, in the northern. For the latter, Völk *et al.* estimated — as a conservative minimum — a VHE flux of $F_\gamma(\geq E) = 4.7 \times 10^{-13}$ cm^{-2} s^{-1}, in the TeV range [137]. Despite being a conservative estimate, this value is well within the capabilities of modern Cherenkov detectors (see e.g. Fig. 5.9). However, the only detection of such object so far has been that of NGC 253 by the CANGAROO collaboration (> 10-σ level), who reported a flux of $\approx 1.4 \times 10^{-11}$ cm^{-2} s^{-1} above 400 GeV, corresponding to $\approx 15\%$ of the flux from the Crab nebula [138],[139]. The emission was seen as coming from an extended source (0.3–0.6°) that was somewhat wider than the optical size of the starburst galaxy ($\approx 28' \times 7'$). As a result, it was interpreted as inverse Compton interaction in an extended halo of multi-TeV CR

Source	Possible Counterpart	Class	T (h)	σ	$F(>200\ \mathrm{GeV})$ ($10^{-12}\ \mathrm{cm}^{-2}\ \mathrm{s}^{-1}$)	ν
HESS J1616−508	PSR J1617−5055	PWN	10.2	16.3	43.3 ± 2.0	2.35 ± 0.06
HESS J1632−478	IGR J16320−4751	XRB	4.5	4.5	28.7 ± 5.3	2.12 ± 0.20
HESS J1634−472	IGR J16358−4726/G337.2+0.1	XRB/SNR	6.6	4.2	13.4 ± 2.6	2.38 ± 0.27
HESS J1640−465	G338.3−0.0/3EG J1639−4702	SNR/UID	14.3	13.6	20.9 ± 2.2	2.42 ± 0.15
HESS J1713−381	G348.7+0.3	SNR	4	13	4.2 ± 1.5	2.27 ± 0.48
HESS J1745−303	3EG J1744−311	UID	35.3	4	11.2 ± 4.0	1.82 ± 0.29
HESS J1804−216	G8.7−0.1/PSR J1803−2137	SNR/PWN	15.7	13	53.2 ± 2.0	2.72 ± 0.06
HESS J1813−178	G12.82−0.02	SNR	9.7	13.5	14.2 ± 1.1	2.09 ± 0.08
HESS J1825−137	PSR J1826−1334/3EG J1826−1302	PWN/UID	8.4	8.1	39.4 ± 2.2	2.46 ± 0.08
HESS J1834−087	G23.3−0.3	SNR	7.3	8.1	18.7 ± 2.0	2.45 ± 0.16
HESS J1837−069	AX J1838.0−0655	UID	7.6	12.2	30.4 ± 1.6	2.27 ± 0.06

Table 1.6: (from [69]) Summary of the newly discovered sources in the H.E.S.S. survey, for which there exist possible associations (column 2). Where more than one counterpart has been considered, a slash has been inserted between the candidate sources. The object class of the possible counterpart is shown in the third column: "PWN" stands for "pulsar wind nebula"; "XRB", for "X-ray binary"; "SNR", for "supernova remnant"; "UID", for "unidentified source" in the EGRET and *ASCA* maps; and "BH" stands for "black hole candidate". The fourth column presents the exposure times (T) of the data sets that were used in the respective analyses (livetime); and the fifth, the resulting statistical significance of the signal in standard deviations above the background. The sixth column shows the integral fluxes, $F(>200$ GeV), above 200 GeV. Finally, the seventh column shows the spectral index of the differential power-law spectrum as was fitted across the investigated range. The errors are statistical only.

Source	σ	E_{th} (TeV)	$F(>E_{\mathrm{th}})$ ($\mathrm{cm}^{-2}\ \mathrm{s}^{-1}$)	Reference
HESS J1614−518	10.1	0.2	$5.8 \pm 7.7 \times 10^{-11}$	[69]
HESS J1702−420	4	0.2	$1.6 \pm 1.8 \times 10^{-11}$	[69]
HESS J1708−410	6.8	0.2	$8.8 \pm 0.7 \times 10^{-12}$	[69]
HESS J1303−631	21	0.38	$1.5 \pm 0.3 \times 10^{-11}$	[133]
TeV J2032+4130 (Whipple)	3.3	0.6	$5.6 \pm 1.6 \pm 0.8 \times 10^{-12}$	[135]
TeV J2032+4130 (H.E.G.R.A.)	7	1	$5.9 \pm 3.1 \times 10^{-13}$	[134]

Table 1.7: Summary of the unidentified VHE sources as of 2005. The second column presents the resulting statistical significance of the detected signal in standard deviations above the background. The fourth column shows the integral flux, $F(>E_{\mathrm{th}})$, above the respective energy threshold, E_{th} (column 3). Where one or more errors are stated, the first corresponds to the statistical uncertainty and the second, to systematic.

electrons around NGC 253. A dark matter annihilation scenario was also discussed as the emission mechanism (see [140],[136]), but the authors concurred that the produced VHE flux in such scenario is too low to be detectable by current Cherenkov experiments.

Given those early reports of VHE γ rays from NGC 253, the H.E.S.S. collaboration decided to observe this source for 38 h in 2003 [136]. After applying the suitable selection criteria, a 28-h data set was analysed, with an estimated energy threshold of 300 GeV. No significant emission was detected, either from a point-like source or an extended region matching the CANGAROO source (0.5°). In the latter case, the derived upper limit was roughly a factor 2 lower than CANGAROO's measurements ($F_{\rm ul}(> 300$ GeV$) = 6.3 \times 10^{-12}$ cm^{-2} s^{-1}), and hence inconsistent with the previously seen emission. A variable-flux scenario as the reason behind the non-detection could be discarded given the extended nature of the source. Also, through various theoretical considerations that were discussed in the relevant publication, the expected flux from NGC 253 was consistent with the derived upper limit. On the other hand, the extended emission that was observed by CANGAROO could not be explained by inverse Compton interactions of cosmic-ray electrons, as it leads to negligible fluxes and harder spectra compared to CANGAROO's.

Active Galactic Nuclei (AGN)

Amongst the extragalactic objects that emit VHE γ rays there are Active Galactic Nuclei, or AGN. Although their nature was a matter of debate in the days of their first discovery, the present consensus describes them as active galaxies whose core is many times more luminous than the cores of normal galaxies and harbours sites of relativistic particle acceleration. According to the "standard model" (or "unified theory") of AGN, it is thought that their core comprises a $10^{8\text{-}9} M_\odot$ super-massive black hole, the size of the solar system, that accretes galactic material [1]. Due to the violent nature of the accretion process in AGN, some of them exhibit plasma jets which are aligned with the black hole's axis. Surrounding the accretion disk, which is a few parsecs across, there is a torus of dense dust that is heated up by the radiation from the accretion process. There exists evidence that the jet emission propagates perpendicular to the plane of the torus, although conclusive arguments that this is the case for all AGN have not been brought to light, yet. In addition to the torus, the standard model states that the central compact source is surrounded by a random distribution of gas and dust clouds that are also heated by the accretion's

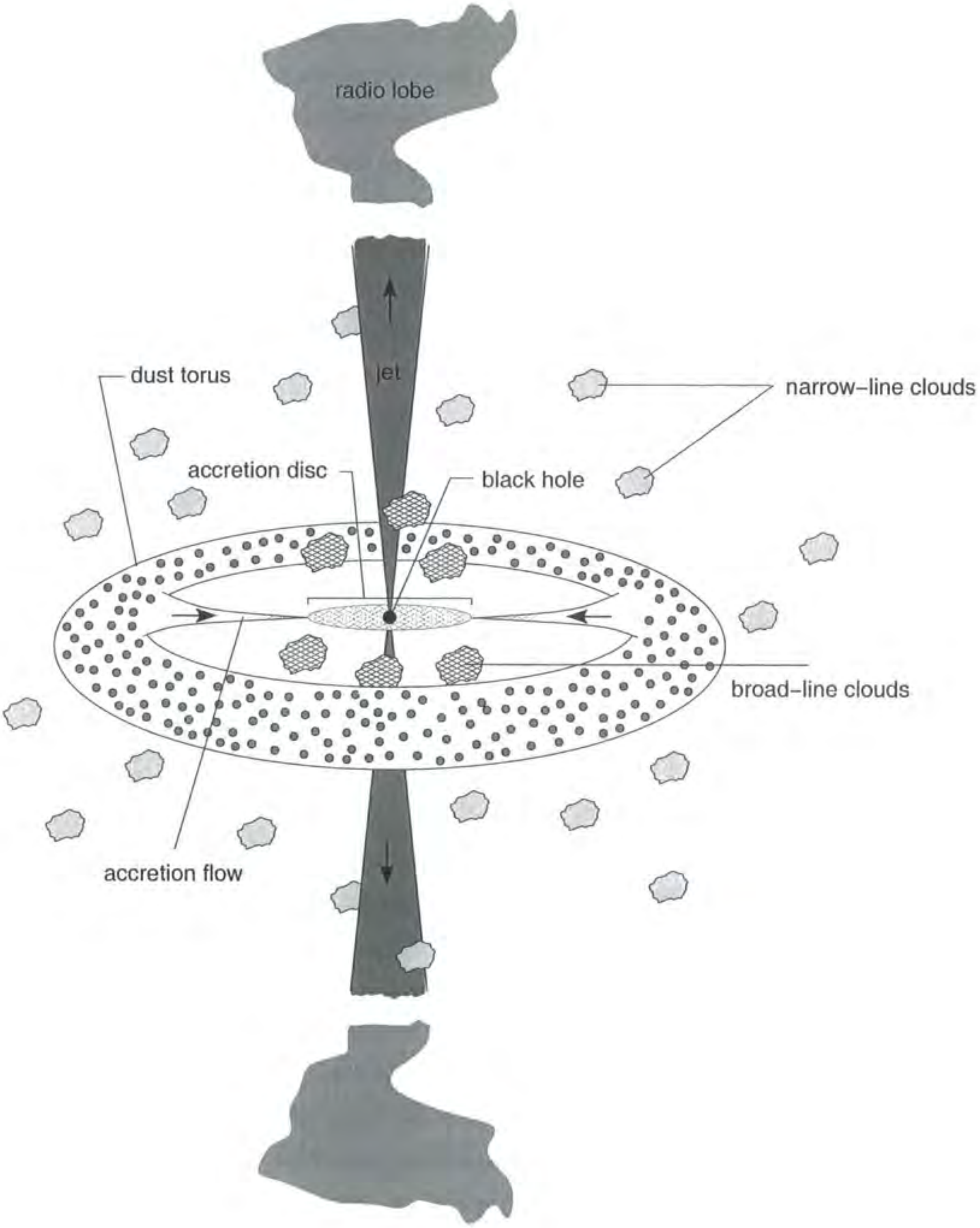


Figure 1.14: The main components of the "standard model" of AGN. The figure is not to scale.

radiation. Finally, far from the AGN's core, at hundreds of kiloparsecs distance, the jets dissipate their energy into the characteristic radio lobes observed in many quasars. A graphical representation of the standard model's central components is presented in Fig. 1.14.

The earliest classification scheme of AGN was based on the observed optical spectra, and more specifically the presence of narrow emission lines in the AGN's spectrum. In the light of the standard model, these lines exist due to the randomly distributed population of dust clouds around the active galaxy. On the other hand, the absence of broadband emission from some AGN, from optical to X-rays, has been explained in terms of thick dust clouds obscuring the direct emission from the central source; in those cases, radio, infrared and γ-ray radiation can still be visible, and narrow emission lines from the outer gas clouds may still be present.

The different appearance of AGN has also been explained based on the standard model. For instance, if the active galaxy is viewed edge-on, emission from the AGN's core can be completely obscured by the dust torus; then the AGN is seen as a *radio galaxy* with emission lines clearly visible in the spectrum and the luminosity dominated by the lobes' radio emission. However, of particular interest for VHE astronomy are AGN whose jets are roughly aligned with our line-of-sight. In those cases, the core becomes more apparent and part of the jet's emission can be seen; dust-cloud emission can be visible too. As the alignment becomes better, the jet's emission is now the dominant feature of the AGN's spectrum, and the object is called a *blazar*. Finally, when the relative alignment allows us to view the emission directly from the base of the jet, the AGN is called an *extreme* blazar.

Blazars

Blazars can be subdivided into two major classes according to their spectral features: those whose spectra exhibit clearly visible emission lines are called Flat Spectrum Radio Quasars (FSRQ); otherwise, if the emission lines are weak or absent, the blazar is called a BL Lac object, following the name of the first object discovered with such property.

Ground observations of distant BL Lac are the key into probing the flux of the Extragalactic Background Light (EBL). The EBL flux observable at the present date is the sum of all the starlight contributions from all the galaxies throughout the evolution of the universe; light which was emitted during the earliest star formation may also have an important contribution.

Of particular interest for VHE astronomy is the IR part of the EBL spectrum. Over cosmological distances ($z \gg 0.01$), there is a significant probability of absorption of TeV γ rays by IR EBL photons via the pair-production process (see section 2.7). Especially, observations at large zenith angles utilise the increased effective areas of the detector (see e.g [141]) and allow the probing of the multi-TeV spectra of AGN, which are expected to be affected by the EBL. Further discussion about EBL measurements with VHE experiments takes place later in this section (see page 71).

Despite their cosmological value, VHE detections of blazars had to wait until these objects had already been established as γ-ray sources in the MeV–GeV range [14]. Interest turned to AGN after EGRET's discoveries in the HE range, and it was not long before the first BL Lac object, Mkn 421, was detected with the Whipple telescope above 500 GeV [142]. In addition, it is worth mentioning Mkn 501, which also joined the list of TeV BL Lac, after the observations of Whipple and H.E.G.R.A. in the mid-1990s [143],[144]: the former resulted in the detection of a few % of the Crab nebula's flux from that source, whereas H.E.G.R.A. observed Mkn 501 during a period of unexpectedly high activity (1997), which resulted in flux peaks of 10 Crab (the mean flux was 3 Crab).

However, the Whipple collaboration detected another two BL Lac objects in the period 1999–2002: **1ES 1959+650** and **H 1426+428**. The 2002 data from 1ES 1959+650 showed a > 20-σ significance above $\sim$ 600 GeV, in 39 h of on-source observations [145]. In addition, H 1426+428 was detected in 2000–2001 observations at a significance of 5.5 σ, above 280 GeV [146]. Both sources had been observed before at lower energies — by experiments like *BeppoSAX* — but above 100 MeV EGRET did not detect HE emission from their direction.

However, the lowest-energy observations of 1ES 1959+650 have been with the MAGIC telescope [147]. MAGIC observed this source for 6 h in September and October 2004, and despite being in a state of low optical and X-ray activity during that period, 1ES 1959+650 was detected above 180 GeV at an $\approx$ 8-σ level.

In 2000–2003, H.E.S.S. probed a selected list of AGN, whose distance spanned from $z = 0.00183$ to $z = 0.333$ [148]. The zenith angles of observation were in the Z.A.=11–64° range, and an exposure time ranging from $\sim$ 1–8 hours was given to the individual observations. Apart from **Mkn 501**, which produced a 3.1-σ signal, the rest were undetected; but the most constraining upper limits to that date were derived for those other AGN: these ranged from 0.4–5.1% of the Crab nebula's flux. For Mkn 501 — which was observed at the large Z.A. of 64° — an estimated flux of 15% of that from the Crab nebula above 1.65 TeV was derived.

A nearby and well-studied AGN is the first to have been detected in VHE γ rays, **Mkn 421**. It has a redshift of z = 0.031 and exerts a highly variable emission on an hourly–daily time-scale. Its average γ-ray flux was seen to increase tenfold in the 1994 Whipple observations (from 15% to 150% of the Crab nebula's flux), over a 2-d period (see Fig. 1.15) [149]. Despite these variations, Mkn 421 is now considered a standard γ-ray source. In 2001, it exhibited an extraordinary 3-month period of activity, during which it outshone the γ-ray signal from the Crab nebula (see e.g. [150]). During this outburst, a significant amount of γ rays, at the rate of $\approx$ 7.7 ph min^{-1} above 140 GeV, was also collected with STACEE [151]. This allowed the recording of Mkn 421's lightcurve between 50 and 500 GeV. Follow-up observations with H.E.S.S., in April and May 2004, resulted in the construction of the overall spectrum of Mkn 421 from 1 to 50 TeV, which showed a clear curvature: the 1–15 TeV range was well-sampled and could be fitted with a simple power law (ν = 2.1), with an exponential cut-off at 3.1 TeV; and at higher energies, in the 30–50 TeV range, an upper limit was calculated.

As the detectability of this object is no longer an issue (e.g. the 2004 H.E.S.S. data led to a significance of 114 σ [152]) interest has turned to Mkn 421's variability. In the 2004 multi-wavelength campaign on Mkn 421, the VERITAS collaboration and the *RXTE* X-ray satellite conducted contemporaneous observations of the object, in order to correlate its variability and spectral characteristics in keV and TeV energies, respectively [153]. H.E.S.S. observations were carried out almost simultaneously in that period, too. Although these observations did not occur simultaneously with those of *RXTE*, an effort was made to combine the H.E.S.S. data coverage with that of the other experiments. A combination of *RXTE*, Whipple and H.E.S.S. observations allowed the study of flux variability correlation between X-rays and GeV γ rays (> 400 GeV), and TeV γ rays (> 2 TeV). It was found that the RMS width of the diurnal variability in the TeV observations was roughly twice that in the sub-TeV range. More specifically, Mkn 421's TeV flux was found to vary by up to a factor of 4.3 on a daily scale, and for nights of high flux, an hourly variability was observed. Above 2 TeV, the average integral flux was 3 Crab, over the data span, and peaked at a diurnal average of 5 Crab during nights of high flux. Moreover, an interesting feature of Mkn 421's emission was the spectral hardening with increasing flux: it showed that periods of increased activity from this AGN are accompanied by an enhanced multi-TeV component. Also, the cut-off energies were found to vary across 1.5–3.5 TeV [152].

The spectral characteristics of Mkn 421 are very similar to Mkn 501: both can be fitted with a power law with an exponential cut-off at a few TeV. (In 1999, H.E.G.R.A. reported fits to Mkn 501, with a cut-off at ≈ 5.5 TeV [154].) Moreover, both objects are at a relatively short distance: $z = 0.031$ for Mkn 421, and $z = 0.034$ for Mkn 501. This led to the conclusion that the spectral curvature of these AGN must be an intrinsic characteristic rather than due to EBL absorption.

The brightest southern-sky BL Lac in X-rays, and the brightest of its kind in UV, is the High-Frequency BL Lac (HBL) **PKS 2155−304**: it is a BL Lac whose energy spectrum peaks in the UV–X-ray range. Its high variability permitted only one detection, in GeV–TeV energies (with the Durham Mark 6 telescope), in the period 1997–2002. The reported flux was $F(> 300 \text{ GeV}) = 4.2 \times 10^{-11} \text{ cm}^{-2} \text{ s}^{-1}$ [155].

A second and more decisive TeV detection of PKS 2155−304, at the significance level of 44.9 σ, came in the 2002–2003 observations with H.E.S.S. [156]. Temporal analysis showed that the flux varies on hourly–monthly scale, whereas the fitted power-law spectrum does not change over time. Despite the evident variability, the estimated flux above 300 GeV from those new observations was in reasonable agreement with Durham's detection. The different fitting approaches that were tried above 300 GeV included simple and broken power-law fits, as well as a power-law fit with an exponential cut-off. However, it was concluded that the latter was not significantly better — resulting in a better fit to only 85% probability. Hence, despite the desirable scenario of having a spectrum with features — which could imply EBL absorption — this case could not be verified with confidence; and the authors argued that even in such case, one cannot be confident that this is due to the an EBL component, but it could also be an intrinsic spectral feature of this particular AGN.

Another result that is of great value to the study of VHE AGN spectra was the detection (at the 6.7-σ level) of the BL Lac **PKS 2005−489** in H.E.S.S. observations in 2004 [157]. Not only was it the first time the PKS 2005−489 was being detected in the VHE regime, but it also became only the second such AGN known in the southern hemisphere. Interestingly, H.E.S.S. did not detect it in observations that took place in 2003, and the derived upper limit ($F_{\rm ul}(> 200$ GeV$) = 5.2 \times 10^{-12}$ cm^{-2} s^{-1}) was lower than the 2004 flux ($F(> 200$ GeV$) = 6.9 \pm 1.0 \pm 1.4 \times 10^{-12}$ cm^{-2} s$^{-1} \sim 0.025$ Crab), which indicated a variable emission on a yearly scale; but a temporal investigation for a shorter-term variability showed a stable flux. This conclusion was strengthened by parallel X-ray observations in 2003–2004, which revealed a threefold increase of the average X-ray count rate over 2004. A comparison with older X-ray data, however, showed that even the 2004 activity was much lower than that in 1998, which led to the conclusion that H.E.S.S. detected PKS 2005−489 during its "quiet" state. Furthermore, this BL Lac's spectrum was the softest ever measured ($\nu = 4$) in its class, but this is unlikely to be due to EBL absorption given its proximity ($z = 0.071$).

Finally an AGN that was recently detected with H.E.S.S. is the extreme blazar **H 2356−309**. Lying at $z = 0.165$, H 2356−309 has been long thought as a strong candidate for TeV emission [158],[159]. A 40-h exposure on this source between June and December 2004 yielded 10 σ total excess [160]. The observed integral flux above 200 GeV showed variations on monthly timescales, but no variability was detected on shorter scales. The maximum activity occurred in September, when the signal excess was 6.7 standard deviations above the background.

Radio Galaxies

Although blazars dominate the class of VHE emitters, "radio galaxies" are also targets for VHE observations. Their name was given as a good description of their intense radio luminosity, which can exceed that of normal galaxies by three orders of magnitude. However, it is now understood that they are, like the BL Lac, Active Galactic Nuclei but much closer to us — the closest being

Centaurus A (Cen A), with $z = 0.0018$. Radio galaxies exhibit lobes of radio emission that stretch out far beyond the visible boundaries of the host galaxy; and at the base of of the lobes there exist plasma jets that appear to connect the lobes to the centre of the galaxy [1]. The various theoretical models that predict VHE emission from this type of AGN include inverse Compton interaction from relativistic leptons, TeV production in the emanating plasma jets, as well as various hadronic models; and the scenario of neutralino annihilation is also considered a possibility. It is clear that these objects are sites of particle acceleration, and for that reason they are considered prime candidates for the source of the most energetic cosmic rays.

Naturally, the proximity of radio galaxies, as well as the fact that they are sites of particle acceleration, has triggered a high interest in VHE observations. In particular, they have recently become the centre of attention with the observations of the giant radio galaxy **M 87**, with the H.E.G.R.A. and H.E.S.S. experiments. But the first detection of a signal from a radio galaxy came from the observations of Cen A by Grindlay *et al.* in Australia [161]. Grindlay and his colleagues used the atmospheric Cherenkov technique to detect a 4.5-σ signal from the direction of Cen A. The integral flux above 300 GeV was $4.4 \pm 1.0 \times 10^{-11}$ cm^{-2} s^{-1}. However, subsequent observations did not confirm this result: the Durham group, operating the Mark 3 detector, reported a 3-σ upper limit above 300 GeV (7.8×10^{-11} cm^{-2} s^{-1}) from 44-h observations during 1987–1988 [162]; observations made in 1995 with CANGAROO-I placed the flux above 1.5 TeV below 1.28×10^{-11} cm^{-2} s^{-1} [163]; finally, in 1997, a re-investigation of Cen A by Durham, this time with Mark 6, resulted again to a 3-σ upper limit equal to $F(> 300\ \text{GeV}) = 5.2 \times 10^{-11}$ cm^{-2} s^{-1} [164].

During the periods 2003 and 2004–2005, VHE emission from M 87 was detected with H.E.S.S., at the level of $\approx 6\sigma$; the integral flux above 730 GeV was estimated to $\sim 1\%$ that from the Crab nebula [165]. Previous observations of M 87, in 1998–1999 with H.E.G.R.A., had detected (at the 4.7-σ level) a flux above the same energy threshold equal to 3.3% that of the Crab nebula's [166]. Hence, it was concluded that the TeV emission exhibits a degree of variability, although more observations are needed to establish this.

The short list of 11 AGN recently observed in the VHE range is presented in Ta-

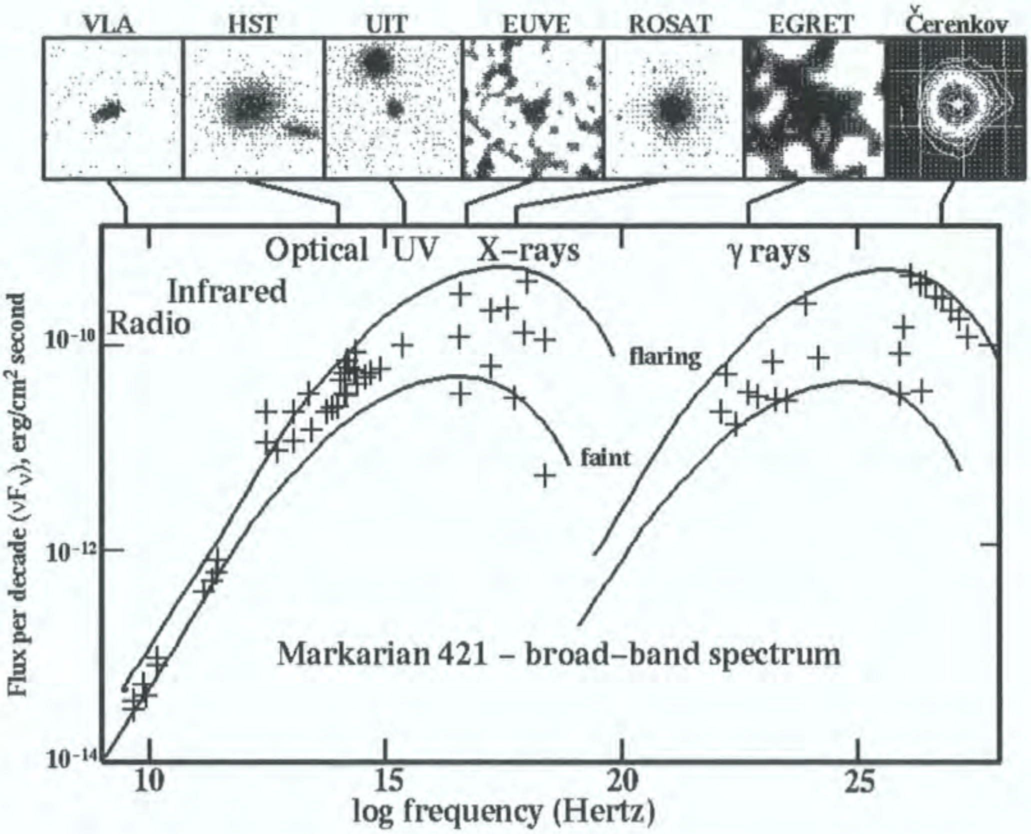


Figure 1.15: (from [169]) Mkn 421 across the electromagnetic spectrum. In its faint state (lower right curve), this AGN can only be detected with modern, ground-based γ-ray detectors. The experimental data for the γ-ray spectrum, shown with crosses, are taken with the EGRET instrument and the Whipple Cherenkov telescope, for energies spanning from MeV to GeV and hundreds of GeV to TeV, respectively.

ble 1.8. Amongst others it includes the extreme blazar **1ES 1101−232**, which is the most distant object detected at TeV energies so far, with $z = 0.186$ [167]; although there are strong indications that another BL Lac, BL Lac **PG 1553+113**, which was detected with H.E.S.S. in 2005, is even further, at $z > 0.25$ [168].

γ-ray Bursts (GRBs)

As the mystery of γ-ray bursts (GRBs) unfolds, more sources are expected to be associated with these violent events. Unfortunately, VHE experiments are not optimal for GRB research. Except for Milagro, a water-pond detector that is capable of monitoring the whole overhead sky, all other Cherenkov telescopes have narrow

Source	Class	z	T (h)	σ	$E_{\rm th}$ (TeV)	$F(> E_{\rm th})$ ($\rm cm^{-2}\ s^{-1}$)	Reference
PKS 2005−489	HBL	0.071	24.2	6.7	0.2	$6.9 \pm 1.0 \pm 1.4 \times 10^{-12}$	[157]
PKS 2155−304	HBL	0.117	≈ 63	44.9	0.3	(1.18 ± 0.15)-$(7.77 \pm 1.63) \times 10^{-11}$	[156]
1ES 1101−232	HBL	0.186	43	≈ 12	0.16	$\approx 7.4 \times 10^{-12}$	[167],[170]
1ES 1959+650	HBL	0.047	6	8	0.18	$3.73 \pm 0.41 \pm 0.35 \times 10^{-11}$	[147]
PG 1553+113	HBL	< 0.74	7.6	> 5	0.2	$4.8 \pm 1.3 \pm 1.0 \times 10^{-12}$	[168]
H 2356−309	extreme BL Lac	0.165	39	9	0.2	$4.1 \pm 1.4 \times 10^{-12}$	[160]
H 1426+428	extreme BL Lac	0.129	31	5.5	0.28	$2.04 \pm 0.35 \times 10^{-11}$	[146]
1ES 2344+514	extreme BL Lac	0.04	53	3	2	$< 4.2 \times 10^{-12}$	[171]
Mkn 421	BL Lac	0.031	14.7	114	0.2	$\approx 1.8 \times 10^{-11}$	[152]
Mkn 501	BL Lac	0.034	1.8	3.1	1.65	$1.5 \pm 0.6 \pm 0.3 \times 10^{-12}$	[148]
NGC 253	SB	0.0008	28	2	0.3	$< 6.3 \times 10^{-12}$	[136]
M 87	radio galaxy	0.004	45	5.8	0.73	$\approx 3.1 \times 10^{-13}$	[165]

Table 1.8: Summary of recent results from observations of AGN. The object class is presented in the second column using the following abbreviations: "HBL" stands for "high-frequency BL Lac", a BL Lac whose energy spectrum peaks in the UV–X-ray range, and "SB", for "starburst galaxy". The third column shows the redshifts, z, of the respective objects; the redshift shown for PG 1553+113 is an upper limit based on EBL absorption. The fourth column presents the exposure times, T, that were used in the respective analyses (livetime); and the fourth, the resulting statistical significance of the signal in standard deviations above the background. Where only an upper limit is reported (indicated with the inequality sign under column 7), the significance shown corresponds to the confidence level of that upper limit. Finally, the seventh column shows the integral fluxes, $F(> E_{\rm th})$, above the respective energy thresholds, $E_{\rm th}$ (column 6): where one or more errors are stated, the first corresponds to the statistical uncertainty and the second, to the systematic.

fields of view — typically a few degrees — which makes it unlikely that a GRB will occur in their window. Furthermore, their slewing speeds are much slower than those of satellite telescopes. There exists one claim that a GRB *has* been observed from the ground while the phenomenon was unfolding. During 1997–1998 observations, Milagrito — the prototype of Milagro — detected an $\approx$ 4-σ signal excess that was spatially and temporally correlated with one of the 54 GRBs observed with BATSE that period [172]: the 7.9-s GRB, **GRB 970417a**. The chance probability of having detected at least one such signal in a sample of 54 GRBs was 1.5×10^{-3} ($\approx$ 3 σ). The detection threshold was 500 GeV, and parallel observations with BATSE (the Burst and Transient Source Experiment) did not result in a positional coincidence, as the GRB appeared weak in BATSE's energy range.

Nevertheless, the information to be gained from VHE observations is very valuable for theoretical modelling, and of particular interest is the determination of a cut-off in the GRB spectra. Due to the cosmological distances involved, a spectral steepening around 100 GeV is expected. This steepening is caused by the pair-production absorption of γ rays on the low-energy photons of the IR background [173],[174].

Extragalactic Background Light (EBL)

The spectrum of the Extragalactic Background Light has always been a focal point for cosmology because it directly linked to early star formation. Direct measurements of the diffuse, infrared EBL have already been conducted with experiments like the Far Infrared Absolute Spectrophotometer (FIRAS) or the Diffuse Infrared Background Experiment (DIRBE) [175],[176],[177]. Unfortunately, these measurements suffer from contamination by local sources of IR emission.

Alternatively, VHE γ-ray observations are capable of *in*direct measurements of the EBL. The transmission of HE and VHE γ rays over cosmological distances is limited by pair-production interactions on optical and IR photons, respectively. TeV photons are expected to be absorbed by IR radiation produced during stellar formation. Moreover, because the cross section of the γ–γ interaction has a strong peak at the characteristic energy at which the γ rays interact with the IR photons, the EBL spectrum is expected to imprint some of its features on the observed γ-ray one.

The most suitable γ-ray spectra for inferring the EBL density are those of AGN. They provide a measurement that is free of biases caused by the local, Galactic IR emission. However, even AGN spectra are affected by absorption processes, which are unrelated with the IR photon density. Strong radiation fields at the sites of γ-ray production can also cause absorption that can be confused with that of the genuine IR component. To make matters worse, an accurate measurement of the IR flux density requires an equally accurate knowledge of the intrinsic AGN spectrum, which is usually unknown.

Nevertheless, the special cases of Mkn 421 and Mkn 501, whose spectra are well measured, have allowed us to set limits on the EBL density over a wide range of wavelengths. The results are not constraining with regards to galaxy formation, because the upper limits are well above the expected IR flux from normal galaxy formation. However, they do rule out several exotic mechanisms of IR emission [178],[179]. Interestingly, both of the above AGN exhibit a spectral cut-off in the 3–6 TeV range, and given the fact that they have equal redshifts their cut-offs could be attributed to a pair production mechanism [180],[181]. On the other hand, this common feature could also be a coincidence or an intrinsic characteristic of AGN emission [182]. A definite conclusion could be drawn from measurements of other similar AGN, like H 2356−309 and 1ES 1101+232. Although the spectra for these two sources are less accurate, their larger distances compared to the ones above could be the key factor towards the determination of EBL's role in γ-ray transmission. Indeed, the spectrum of H 2356−309 becomes steep above 1 TeV, which may be an indication of EBL absorption. If so, this implies an intrinsic source luminosity which is $\sim$ 2–3 orders of magnitude higher than that suggested by the measurements [160],[170],[1].

Chapter 2

Astrophysical Mechanisms of γ-ray Production

2.1 Introduction

The universe contains a plethora of astrophysical sites of emission, whose properties are widely diverse. For example, there are situations where the magnetic fields can be as weak as 0.1 μG (typical for intergalactic fields), and others where they can be as strong as 10^{12} G or more (a common value for young pulsars) [183],[184]. Hence, it is important to understand the emission mechanisms arising from the behaviour of matter subjected to such magnetic fields. The following sections describe the emission mechanisms most relevant to VHE γ-ray production, and particularly those involved with pulsar VHE emission.

2.2 Synchrotron

According to electromagnetism, a charged particle with velocity $\boldsymbol{v}$, travelling through a magnetic field, $\boldsymbol{B}$, where $\boldsymbol{v} \perp \boldsymbol{B}$, will be forced to move in a circular orbit sustained by the Lorentz force, $\boldsymbol{F} \propto \boldsymbol{v} \times \boldsymbol{B}$. This is a special case of a more probable situation, where the particle has an arbitrary velocity that forms an angle α with $\boldsymbol{B}$, called pitch angle. In such a case, the resultant motion is the synthesis of a uniform motion with velocity $\boldsymbol{v} \cos\alpha$ and a circular one with $\boldsymbol{v} \sin\alpha$: i.e. the helical motion shown in Fig. 2.1. Furthermore, because of the accelerated circular motion, the charged particle will radiate electromagnetic waves. The theoretical treatment of synchrotron emission and the emission from accelerated, charged particles in general is given in various text books: see e.g. [185],[186]. Here, we will only provide a brief overview of the results.

At low energies, the frequency of the radiated EM waves is called Larmor frequency or gyrofrequency and matches the frequency of the particle's circular motion, ω_{g}. The direction of the emission can be derived from the electric field, $\boldsymbol{E}(\boldsymbol{r}, t)$, generated by the moving particle. This field contains two terms: the Coulomb term, $E_{\mathrm{Cl}} \propto 1/R^2$, which does not contribute to the radiation field and can be omitted at large distances, R; and the far-field component, which is proportional to $1/R$. An important approximation, called dipole approximation, can be applied to the retarded potentials involved in the calculation of $\boldsymbol{E}(\boldsymbol{r}, t)$. The resulting angular distribution of the radiated power, $dP/d\Omega = R^2(\boldsymbol{S} \cdot \hat{\boldsymbol{n}})$, is characteristic of dipole emission. $\boldsymbol{S}$ is the Poynting vector, which is proportional to $\boldsymbol{E} \times \boldsymbol{B}$, and $\hat{\boldsymbol{n}}$ is the unit vector in the observer's direction. One implication of this result in the case of synchrotron motion is that most of the radiation is emitted on a plane, normal to the particle's acceleration. Integration over all directions, Ω, of the distribution gives the well-known Larmor formula

$$\frac{dU}{dt} = -\frac{2}{3}\frac{q^2}{4\pi\epsilon_0 c^3}\dot{\boldsymbol{v}}^2 \tag{2.1}$$

As energies go up, approaching relativistic values, the emitted wave properties as seen by a stationary observer have to be transformed from the accelerated reference frame of the particle to the observer's reference frame. In general, this changes the magnitude of the Poynting vector. The dipole distribution changes from axisymmetrical to directional emission which is contained within a cone whose axis matches the particle's direction of motion. As the particle spirals along the helical path, the observer witnesses a precession of the emission's direction around the magnetic field lines. The beamed emission sweeps periodically a section contained between $\alpha - 1/\gamma$ and $\alpha + 1/\gamma$, where α is the pitch angle of the helix, and γ is the Lorentz factor, $(1 - (v \sin\alpha)^2/c^2)^{-1/2}$ (see Fig. 2.1). The pitch angle depends on the ratio between the velocity of the particle's uniform motion and that of the circular motion. Particles that enter a curved magnetic field with velocities almost parallel to the field lines are forced to travel close to the curved field lines, and the pitch angles of their motion are close to zero. Then, the emission is manifested in the forward direction and restricted within a cone of half-angle $1/\gamma$, whose base always remains centred on the B-field line (see Fig. 2.2). This special case of synchrotron radiation is called curvature radiation.

In the relativistic case, it turns out that the Larmor formula for the radiated power has to be boosted by a factor γ^4 [186], and so the synchrotron radiated power per charged particle is given by

$$\frac{dU}{dt} = -\frac{2}{3}\frac{q^2}{4\pi\epsilon_0 c^3}\gamma^4\dot{v}^2 \tag{2.2}$$

This result becomes more interesting if one includes the total energy of the particle, which in this case is conserved, since the work of the magnetic force is zero during the particle's motion. For ultra-relativistic particles, the energy can be approximated with $E \approx pc$. After expressing the acceleration in terms of E and substituting it in Eq. 2.2, we arrive at

$$\frac{dU}{cdt} = -\frac{2}{3}\frac{\mu_0}{4\pi}\left(\frac{qc}{m_0 c^2}\right)^4 B^2 E^2 \tag{2.3}$$

The most important implication of this equation is that heavy particles are poor radiators by comparison. The lighter and more abundant particles, which one would expect to find in astrophysical situations where plasma is accelerated along magnetic field lines, are electrons and positrons. As we will see later, in most theories of high-energy emission, these are the main progenitors of synchro-curvature radiation (see section 3.1.3).

Finally, a very useful knowledge regarding the synchrotron emission is its spectral signature, which is expected to be non-thermal (i.e. different to the Planck distribution). Starting from the relativistic gyrofrequency of the electron, one can calculate the maximum frequency of the emission, ν_{max}. There are mainly two things that have to be considered in order to derive the apparent frequency, observed by a stationary observer, from the gyrofrequency, $\nu_{\text{g}} = eB/2\pi m_0$: the fact that the electron in its circular path almost catches up with the emitted photons, and that the emission is restricted within a narrow cone with $\theta_{1/2} \sim 1/\gamma$. Fig. 2.3 shows an electron in its relativistic, circular orbit, which covers an angular distance equal to $2/\gamma$: i.e. the angular width of the conal emission. Therefore, it shows the positions that the electron takes at the beginning (1) and the end (2) of a single synchrotron cycle, as they are registered by the stationary observer. During the displacement from (1) to (2), the electron emits photons towards the observer's direction. The maximum delay, Δt, between two photons emitted in a single pulse is that between the arrival time of the photons coming from the leading edge of the cone and those

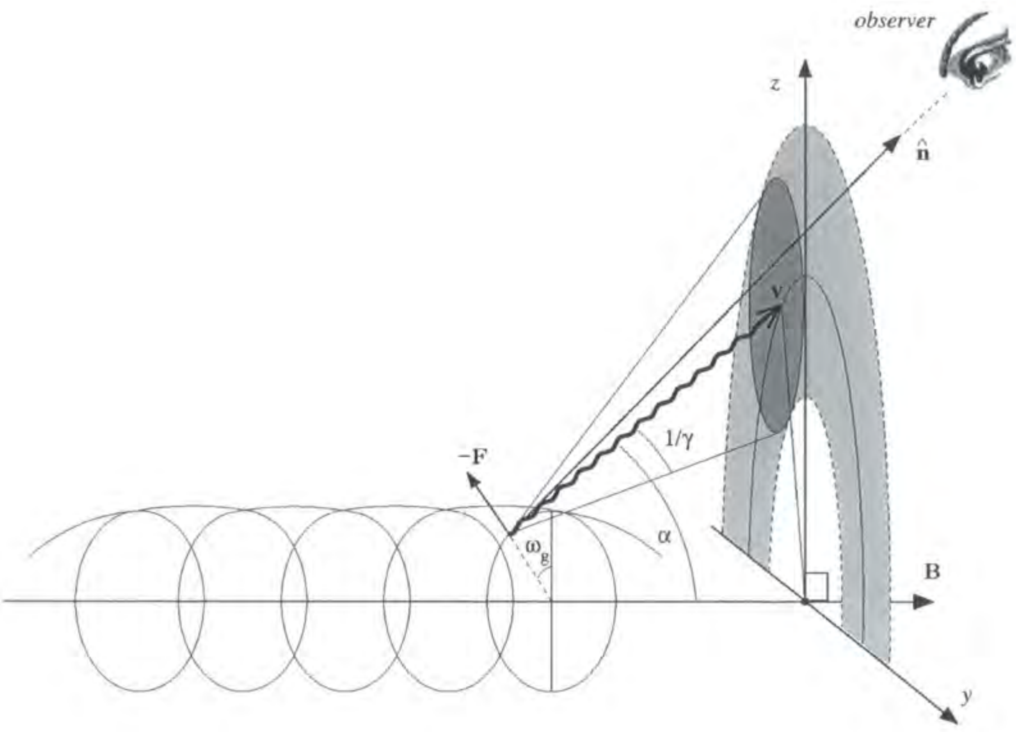


Figure 2.1: The synchrotron emission geometry. The radiation, as seen by a distant observer, is constrained within a conical annulus of half-width equal to the helix pitch angle α and thickness $\sim 1/\gamma$.

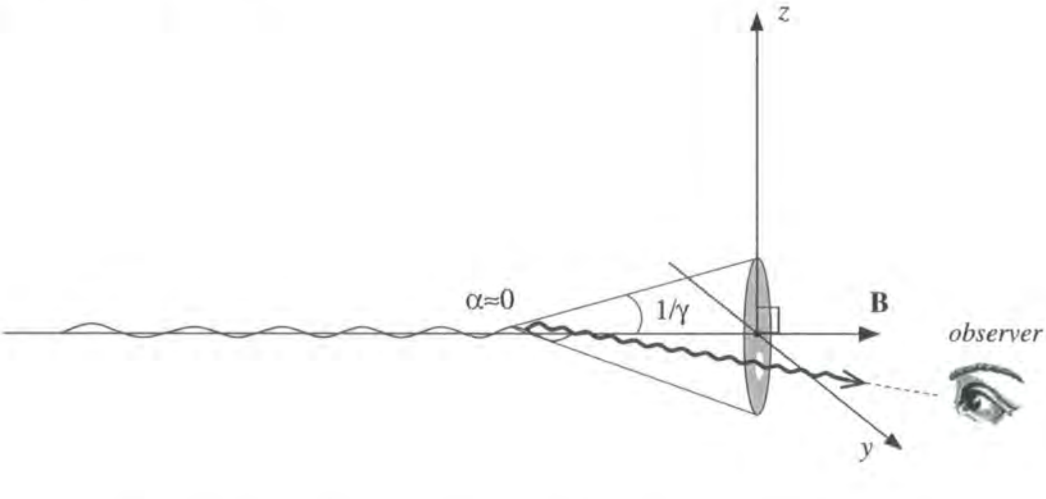


Figure 2.2: The curvature emission geometry. The radiation, as seen by a distant observer, is constrained within a cone of half-width equal to $1/\gamma$.

coming from the trailing edge. So,

$$\Delta t = \frac{L}{v \sin \alpha} - \frac{L}{c} = \frac{L}{v \sin \alpha}\left(1 - \frac{v}{c}\sin \alpha\right) \tag{2.4}$$

For ultra-relativistic revolutions, where $\beta = v \sin \alpha / c \approx 1$, the quantity in the parentheses can be approximated with $1 - \beta \approx (1 - \beta^2)/2 = 1/2\gamma^2$. Furthermore, the term $\Delta t_0 = L/(v \sin \alpha)$ becomes (see Fig. 2.3)

$$\frac{L}{v \sin \alpha} = \frac{r_{\mathrm{g}}\theta}{v \sin \alpha} \approx \frac{1}{\pi \gamma \nu_{\mathrm{g}}^{\mathrm{rel}}} \tag{2.5}$$

where $\nu_{\mathrm{g}}^{\mathrm{rel}} = \nu_{\mathrm{g}}/\gamma$ is the relativistic gyrofrequency. Finally, after substituting with the above approximations, Eq. 2.4 becomes

$$\Delta t = \frac{1}{2\pi \nu_{\mathrm{g}} \gamma^2} \tag{2.6}$$

This result shows that the duration of a synchrotron pulse, as seen by the stationary observer, is roughly $1/\gamma^2$ times shorter than the gyroperiod $T_{\mathrm{g}} = 1/\nu_{\mathrm{g}}$. By decomposing the observed synchrotron pulse into its Fourier components, one finds that the maximum corresponds to frequency $\nu_{\mathrm{max}} \sim \Delta t^{-1}$, or from Eq. 2.6 to $\nu_{\mathrm{max}} \sim \gamma^2 \nu_{\mathrm{g}}$.

Now, the final step is to derive the spectrum of the synchrotron radiation which is emitted by a population of gyrating electrons in a strong magnetic field. Normally, the spectrum is derived by convolving the full frequency spectrum of a single synchrotron-emitting electron with the energy distribution of the population. Here, we will make some assumptions that are not always true, in order to simplify the procedure. First, we assume that the electron energies are distributed according to the power-law function $N(E) = N_0 E^{-s}$, which holds true over a wide range of energies. Secondly, we approximate the electron's full frequency spectrum of emission with a delta function, hence allowing each particle to emit at only one frequency. In that way, the transformation between an electron's energy, E, and its frequency of emission, ν, becomes simple. We conveniently choose that to be ν_{max}, for which we have already derived the relation

$$\nu_{\mathrm{max}} \propto \gamma^2 \nu_{\mathrm{g}} \propto E^2 B \tag{2.7}$$

Based on the stated assumptions, it is now possible to simply derive the population's spectrum following a few proportionalities. The sought-after intensity spectrum describes the directional emission of the synchrotron-emitting source, as mea-

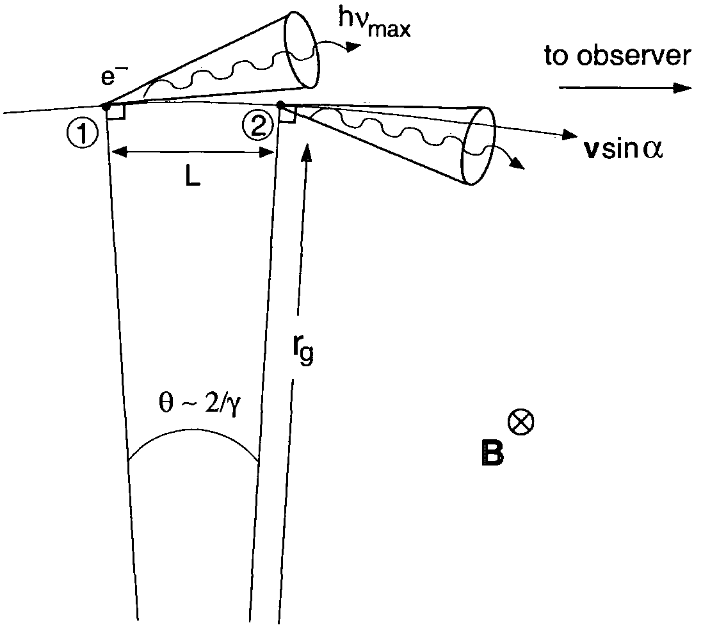


Figure 2.3: A particle spiralling along a magnetic field line emits synchrotron radiation continuously as it gyrates. However, from a fixed position, an external observer sees only a small fraction of the synchrotron duty cycle. That results in the maximum observed synchrotron frequency being roughly γ^2 times higher than the motion's gyrofrequency, ν_g. This figure depicts the fraction of a synchrotron cycle, during which the observer receives the emission. In this figure, the tangential velocity, $v \sin \alpha$, corresponds to the angular velocity of the synchrotron's circular component. Also, during the time interval between positions (1) and (2) which the particle occupies, the observer is swept by the full width, $2/\gamma$, of the conical radiation beam.

sured on Earth, per unit area and per unit time, over a wide range of frequencies (first assumption). We further assume that these frequencies have a 1–1 relation with the energies of the synchrotron-emitting particles. Hence, across an infinitesimal energy interval, $(E, E + dE)$, the radiated energy from the fraction of the population corresponding to this interval will be proportional to the observed intensity, I_ν, in the equivalent frequency interval, $(\nu, \nu + d\nu)$. Therefore,

$$I(\nu)d\nu = -\frac{dU}{dt}N(E)dE \tag{2.8}$$

where the minus sign balances the negative radiation power $dU/dt \propto E^2B^2$ (see Eq. 2.3). Thence, by substituting the aforementioned proportionalities and the power-law function in the RHS of Eq. 2.8, we arrive at

$$I(\nu)d\nu = E^{(2-s)}B^2dE \tag{2.9}$$

Finally, we need to transform the energies to frequencies. For that, we use the proportionality in Eq. 2.7. After some algebra, we conclude to the following expression of the synchrotron spectrum

$$I(\nu)d\nu \propto B^{(s+1)/2}\nu^{-(s-1)/2}d\nu \tag{2.10}$$

The observed spectra of many sources of high-energy emission, like AGN and SNRs, have indices which if used together with the above theoretical prediction result in power-law spectra with $s \approx 2.4$. This number is close to the differential spectral index of cosmic-ray protons with approximately 10 GeV energies [187]. This fact makes synchrotron the possible acceleration mechanism in many astrophysical situations, at least at those energies.

2.3 Bremsstrahlung

Another case of electromagnetic emission due to accelerated charged particles is the Bremsstrahlung radiation. Essentially the same principle which is used for the description of Bremsstrahlung radiated power applies to that of synchrotron emission. However, in the former case the accelerating force is not necessarily perpendicular to the particle's instantaneous velocity. The acceleration/deceleration is typically caused by the electric fields of nuclei lying in the particle's path. This forces the particle to radiate and causes a minor deflection to the particle's direction. In the classical treatment every collision with atoms is accounted for, whereas quantum mechanics calculates a cross section and thus a probability for each collision. Surprisingly, they both agree to the same order of magnitude with regards to the average cross section of many subsequent collisions.

Under the assumption that the deceleration of a particle is due to the field of a nucleus at some distance away, we can use the same generalised expression for the Bremsstrahlung radiation from accelerated particles as for the synchrotron power loss (see Eq. 2.2). However, this time we will assume for simplicity that the relative directions of $\boldsymbol{v}$ and $\dot{\boldsymbol{v}}$ are parallel to each other. The resulting expression is

$$\frac{dU}{cdt} = -\frac{2}{3}\frac{q^2}{4\pi\epsilon_0}\frac{1}{(m_0c^2)^2}\left(\frac{dE}{dx}\right)^2 \tag{2.11}$$

where dE/dx is the accelerating/decelerating force applied to the charged particle [186]. Also, an important feature is that the radiated power is inversely proportional to the square of the particle mass, which means that electrons are strong emitters of Bremsstrahlung radiation, which is also true for synchrotron emission.

Electron Bremsstrahlung is an important process in astrophysical situations where there is a dense gas environment surrounding an electron-emitting source (e.g. SNRs). It can be divided into thermal and non-thermal Bremsstrahlung emission according to the shape of the parent electron distribution. However, in γ-ray astronomy, the electron distributions are typically governed by processes other than thermal. In those cases, the resulting γ-ray spectra depend on these distributions as well as the densities (ρ_g) of the matter with which the electrons interact [1]. The peak energy one can expect from Bremsstrahlung γ rays matches the energy of the most energetic electrons in the distribution.

2.4 Shock Acceleration in SNRs

A very important process in high-energy astrophysics is the acceleration of charged particles in termination shocks. Termination shocks can exist throughout the universe at locations like the regions surrounding a galaxy, where the galactic wind meets the intergalactic matter (IGM), or at the expanding shells of SNRs. It is currently believed that a large part of the cosmic-ray spectrum with energies up to 100 TeV can be attributed to SNRs [188],[189]. For the most energetic cosmic rays — up to $E \sim 10^{19}$ eV — one could look for extragalactic processes. The transportation of cosmic-ray energies beyond $\sim 5 \times 10^{19}$ eV is restricted by the GZK cut-off [190],[191]. However, observations with the cosmic-ray experiment AGASA (Akeno Giant Air Shower Array) concluded that the cosmic ray spectrum extends past the cut-off, and therefore the puzzle of the cosmic-ray origin is still incomplete [192].

During the acceleration phase of a supernova type II, which lasts for a characteristic time $\tau_{\rm acc}$ — until the expanding shell sweeps out its own mass — relativistic charged particles trapped in the irregular interstellar magnetic fields cross the shell's shockfront as a result of their random velocities [193]. In 1949, based on fluid dynamics and the theory of relativity, Enrico Fermi formulated a mechanism which would allow these particles to gain energy form the expanding shell. Although Fermi's first treatment of the problem required large time-scales for the particles to achieve the observed energies, the second version, also known as first order Fermi acceleration, gives a more plausible result in the available $\tau_{\rm acc}$ [194],[195],[196],[197].

The model assumes that, as the supernova shell expands, there is a flow of interstellar matter crossing the shockfront. The matter density behind the shockfront and inside the shell, ρ_1, is higher than that of the interstellar matter ahead of the shockfront, ρ_2. Assuming mass conservation, one can show that the denser the environment through which the matter flows, the slower the flow. In other words

$$\rho_1 v_1 = \rho_2 v_2 \tag{2.12}$$

Fig. 2.4 shows the flow of such interstellar matter through the shockfront, first in the shock's reference frame (R1) and then in the reference frame of the inflowing matter. It is worth noting that the charged particle distribution has isotropically distributed velocities in the frame of reference where no mean drift is observed for the distribution (R2). The shockfront is assumed non-relativistic and that it expands at velocity U. Also, the density ratio ahead and behind the shockfront for a monatomic or fully ionised gas is calculated to be 1/4. We are interested to know what happens when charged particles cross from one side to the other. In Fig. 2.4, case (a) depicts these particles being ahead of the expanding shell. In R1, the interstellar matter has relative velocity U, and after crossing the shockfront it slows down by $3U/4$ (according to Eq. 2.12). Transforming to the reference frame R2, in which the particle distribution has zero drift, one finds that behind the shock there is a constant flow of matter travelling at $3U/4$ towards the shockfront. Therefore, as soon as some of the particles cross on the other side — which happens with probability $2 \sin\theta \cos\theta d\theta$, where θ is the incident angle — they are exposed to the velocity field of the flowing matter. By applying a Lorentz transformation to the flow's moving reference frame in order to calculate the particle energy, and by integrating over all the possible incident angles of the crossing particles, one can prove that each particle receives a boost every time it crosses the shockfront from the region ahead of it to that behind it. On average, it turns out that this energy gain is

$$\left\langle \frac{\Delta E}{E} \right\rangle = \frac{1}{2}\frac{U}{c} \quad \text{per crossing} \tag{2.13}$$

As mentioned, after having crossed the shockfront barrier, the charged particles are exposed to a constant flow of matter. The velocity distribution is again isotropic, this time in the reference frame of the flow *behind* the shock (reference frame R2); see Fig. 2.4b. Again, by transforming all velocities to this frame, we get the situation in which the interstellar matter ahead of the shock flows towards the shockfront at velocity $3U/4$. In other words, the acceleration mechanism as felt by the particles is symmetrical for both directions, and of course since the Lorentz transformation is only dependent on the relative velocity of the reference frames, the energy boost will be the same in both directions. Hence, the particle will be boosted by another $(1/2)U/c$. In total, the particles receive a boost equal to

$$\left\langle \frac{\Delta E}{E} \right\rangle = \frac{U}{c} \quad \text{per cycle} \tag{2.14}$$

A probabilistic treatment based on the classical kinetic theory gives the probability for the particles to escape the acceleration site, and hence one can derive the distribution of the particles that escape [196]. The result is a power-law distribution with spectral index 2: i.e.

$$N(E > E_0) = \text{constant} \times \int_{E_0}^{\infty} E^{-2} dE \tag{2.15}$$

This spectral slope is characteristic of such acceleration sites and is associated with Fermi acceleration [1]. For an observer sitting inside a leaky box model–type region including the source, this spectrum will undergo a steepening. In such a model, the high-energy particles diffuse freely inside the volume that is defined by the region's boundaries and are reflected at the latter. There is a certain probability that some of them may escape from the region and, hence, leave behind a larger population of low-energy particles. The resultant steepening from this process is proportional to

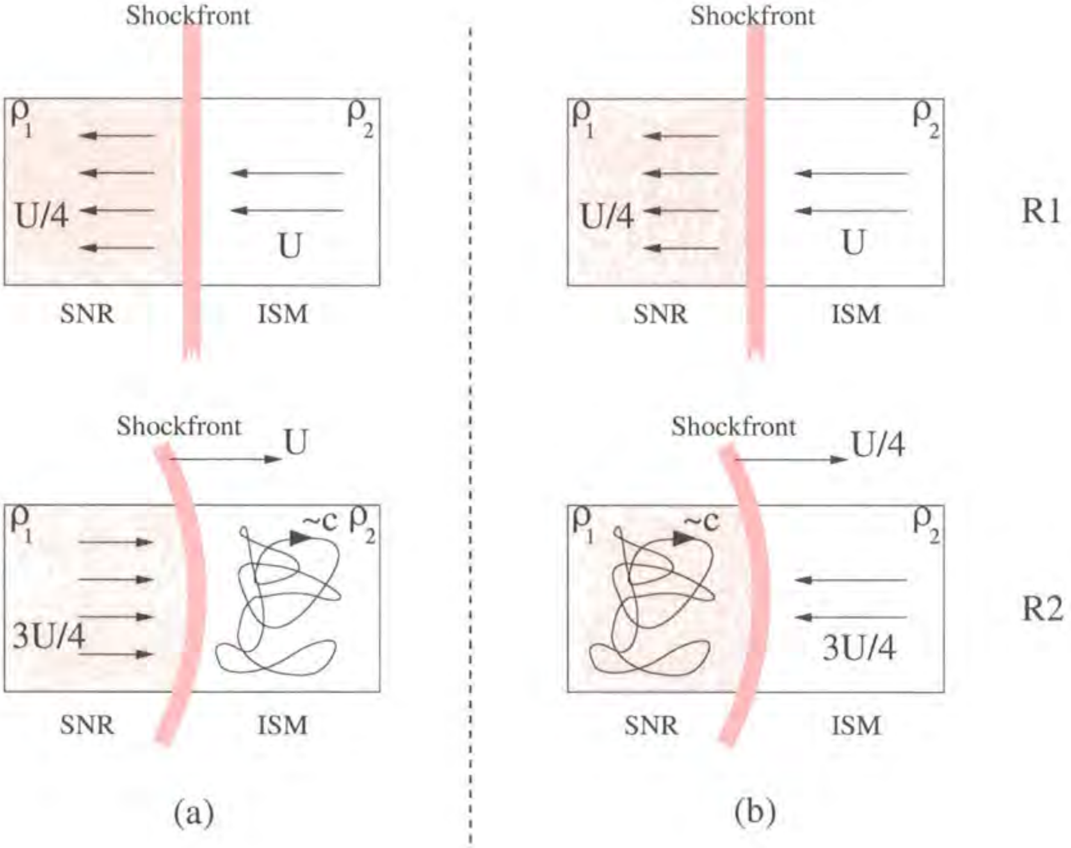


Figure 2.4: The first order Fermi acceleration mechanism. The major assumption is that the distribution of relativistic charged particles on both sides of the expanding supernova shell — with densities ρ_1 and ρ_2 — has isotropically distributed velocities. This enables the particles on both sides to cross the shockfront with equal probability. Case (a) shows the particle distribution being ahead of the expanding shell and, therefore, being part of the interstellar matter which the shock has not yet encompassed. The transformation from the reference frame R1, in which the shock is at rest, to R2, in which the particle distribution has zero drift, shows that after a particle crosses the shockfront, it is exposed to a flow travelling at 3/4 of the shock's velocity. Case (b) shows the particle distribution being behind the expanding shock. After crossing the shockfront, the particles are subjected again to the same velocity field, only in the opposite direction this time. The symmetry suggests an equal boost in both cases, which is $\sim U/c$. Each time a particle crosses the shock front, there is a small probability it will escape from the accelerating region; but particles which remain in the region long enough can be accelerated to relativistic energies. An upper limit on the energy gained during such process has been derived by [189] to be $\sim 10^4$ GeV nucleon^{-1}. Nevertheless, this value is still a matter of debate.

$E^{-0.6}$, and therefore it establishes the shockfront acceleration as a good candidate for the observed cosmic ray spectrum in the local region.

Fermi's first order acceleration mechanism requires the charged particles to diffuse back and forth through the expanding supernova shell. This is clearly a slow process, although faster than the second order acceleration mechanism, which allows for both head-on and follow-up collisions. For a typical SNR in its acceleration process, the shockfront velocity is $\sim 10^3$ km s^{-1}. Therefore, in this phase of the SNR each traversing particle receives about 1% energy gain per crossing.

2.5 Pion Decay

Shock front acceleration provides an explanation for the cosmic ray spectrum up to $\sim$ 100 TeV, but it also provides the energy source for TeV γ rays from SNRs. The interstellar matter (ISM) that surrounds an SNR — mainly atomic hydrogen — is a target for accelerated protons coming out of the SNR shell, and proton–proton collisions are unavoidable. At that point, quantum mechanics tells us that above the kinetic energy threshold of 290 MeV — for the incident protons — collisions between energetic protons and stationary protons will produce pions as well as nucleons. In general, the interaction is of the form

$$p^+ + {}^1\mathrm{H} \overset{E_{\mathrm{kin}}>290\ \mathrm{MeV}}{\longrightarrow} N + N + n_1(\pi^+ + \pi^-) + n_2\pi^0 \tag{2.16}$$

where N denotes a proton or a neutron, and n_1 and n_2 are integers [1].

Almost immediately after the collision (i.e. after characteristic time $\tau_{1/2} \sim 10^{-16}$ s), the produced neutral pions, π^0, decay to two γ rays with $E^0_\gamma \approx 70$ MeV each, where this energy is measured in the pion's rest frame. However, in the laboratory frame, the energy of the γ rays produced in the π^0 decay appears greatly boosted — up to tens of TeV — due to the high Lorentz factors of the pions. (The theoretical properties of the π^0 decay are described in [198], whereas an application to the case of Sgr A East is presented in [199].) For example, a symmetric decay of the π^0 to two photons whose velocities are normal to each other in the pion's rest frame will divide the pion's kinetic energy into two almost equal shares. Therefore, at the energy threshold of the interaction ($\approx$ 290 MeV), the γ rays are produced with $\approx$ 70 MeV energy, and the observed spectrum exhibits a turnover (i.e. a 'bump'). This bump is the signature of γ-ray production via π^0 decay in cosmic-ray sources.

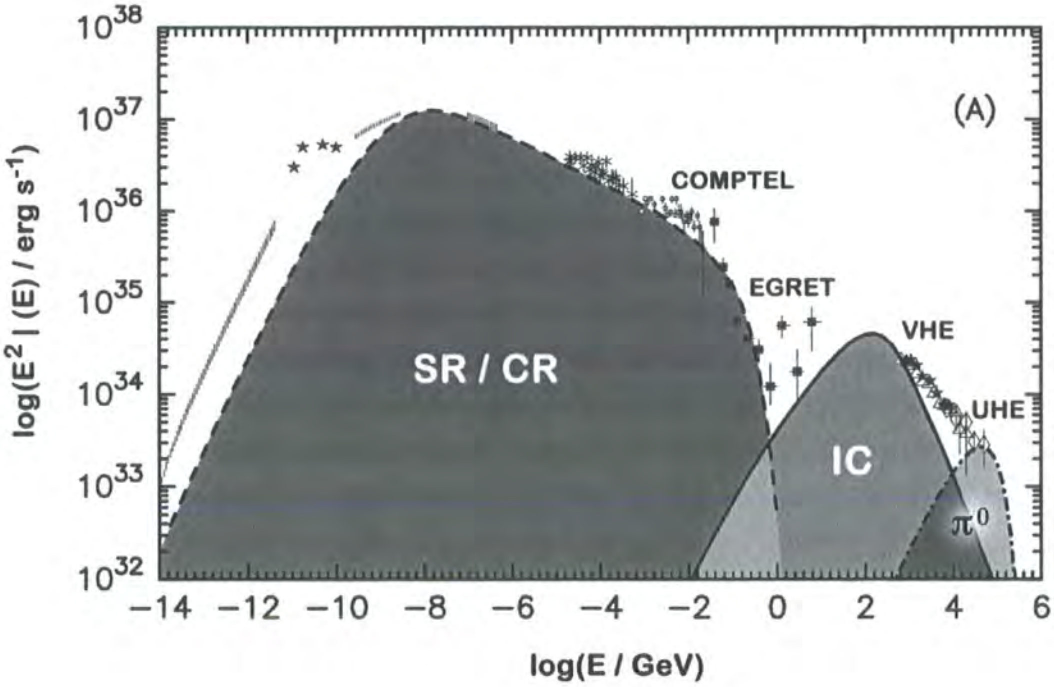


Figure 2.5: (from [200]) The broad spectrum of the Crab pulsar and nebula. Three emission components are clearly distinguishable: a broad synchro-curvature (SR/CR) component extending from low energies up to ≈ 1 GeV, and two high-energy, nebular components due to inverse Compton (IC) and pion decay (π^0) emission.

2.6 Compton and Inverse Compton

Compton scattering as a high-energy emission mechanism is involved in many exotic sources. The principle can be easily reproduced in the laboratory frame at low energies, in which case it is called Thompson scattering. At higher energies, however, the dominant phenomenon is the inverse Compton scattering, described in the following paragraphs. Objects like Pulsars, AGN, etc. are major sources of relativistic particles and, thus, are likely to produce high-energy γ-ray emission via the above mechanisms.

In principle, the Compton or inverse Compton scattering is described by the elastic collision between a photon and an electron with arbitrary kinematic properties. The differentiation between the two processes depends on the properties of both photon and electron. An electron at rest, in the laboratory frame (L), colliding with a photon of energy ϵ, in the same frame, may result in a transfer of energy and momentum from the photon to the electron, with the probability of interaction depending on the cross section of the particular electron–photon pair. In the 1920s, after having observed the inelastic scattering of X-ray photons by aluminium sheets, Arthur Compton derived his famous formula for the energy of the scattered photon [201]. One version of Compton's formula is

$$\epsilon' = \frac{\epsilon}{1 + \frac{\epsilon}{mc^2}(1 - \cos\theta)} \tag{2.17}$$

where ϵ' and ϵ are the photon energies after and before the interaction, respectively, as measured in (L). It is clear from the above that in Compton scattering the deflected photon has energy less or equal to its initial energy, where the case of equal energy corresponds to the Thompson scattering for low-energy (compared to the electron rest mass) photons .

Whether Compton scattering will take place or not for a particular configuration depends on the cross section ($\sigma_{\rm KN}$) of the interaction. The critical factor which $\sigma_{\rm KN}$ depends on is the ratio between the photon and electron energy, $x = \epsilon/(m_e c^2)$. In the laboratory frame (L), the more energetic the photon, the less likely the interaction. However, the impact of the collision is greater for high-energy photons, as the energy of the scattered electron is much increased in those cases. This distinction provides a rough division of the cross-section function into two regimes, according to the value of x. Fig. 2.6 shows the two cross-section regimes, where the probability of interaction can be low or high depending on the energy of the scattered photons, ϵ.

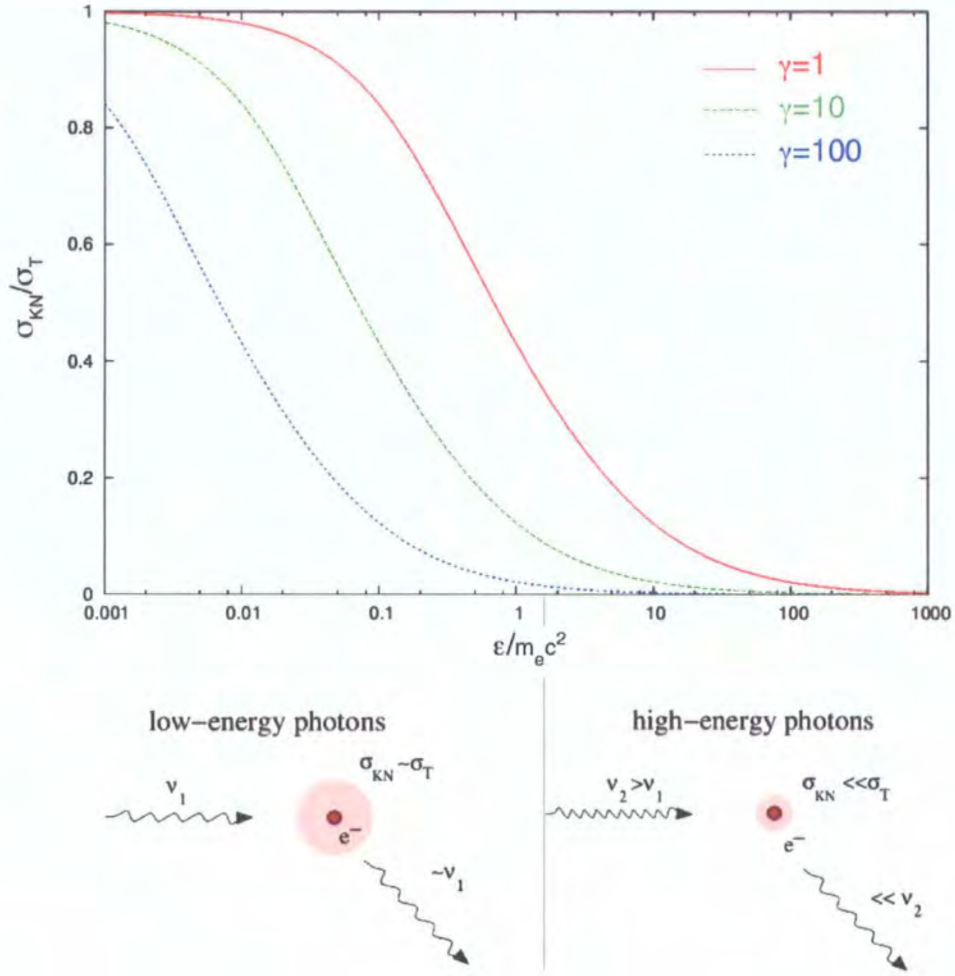


Figure 2.6: The Klein–Nishina cross section variation with incident photon energy. This plot clearly shows the steep decline of the probability for Compton scattering — equivalent to the cross section $\sigma_{\rm KN}$ — with increasing incident photon energy, ϵ. The Klein–Nishina cross section is normalised to the Thompson one, which is the low-energy approximation, and the photon energy is also normalised to the electron rest mass. The maximum probability of occurrence is reached when low-energy photons collide with stationary electrons, i.e. in the Thompson regime, where $\epsilon \ll m_e c^2$. In addition, it can be seen that in the case of moving electrons (Lorentz factor $\gamma > 1$), the Klein–Nishina regime ($\sigma_{\rm KN} \ll \sigma_{\rm T}$) is substantially narrowed. Below the plot, one can see a graphical representation of Compton scattering for the two regimes. The probability of interaction is equivalent to the area that is perpendicular to the photon's direction and encircles the target electron. Scattering occurs if the incident photon crosses this area. However, the more energetic the photon, the smaller the target becomes and, therefore, the less likely the scattering.

The analytical form of the Compton scattering differential cross section was calculated by Klein and Nishina using quantum electrodynamics. When no assumptions are made about the kinematic properties of the interacting photons, its full form is given by

$$\sigma_{\rm KN} = 2\pi r_0^2 \times \left[\frac{1+x}{x^3}\left[\frac{2x(1+x)}{1+2x} - \ln(1+2x)\right] + \frac{\ln(1+2x)}{2x} - \frac{1+3x}{(1+2x)^2}\right] \quad (2.18)$$

where $r_0 = e^2/(m_e c^2)$ is the classical electron radius, and $x = \epsilon/m_e c^2$ is the ratio between the photon energy and the electron rest mass energy, as measured in (L) [198].

A more frequently occurring version of Compton scattering, often met in astrophysical phenomena, is the inverse Compton scattering (ICS). It arises when a fast moving electron, which could be a synchrotron- or curvature-radiation emitter, collides with a low-energy photon (i.e. radio, microwave, IR, etc.). Whenever this interaction occurs, the result is the transfer of energy from the electron to the photon and therefore the upscattering of the latter to higher energies. The whole phenomenon can essentially be described as Compton scattering of the photon in the rest frame of the electron (E). In fact, the probability of interaction is again described by the Klein–Nishina differential cross section. However, the intrinsic difference between the classical Compton scattering and the ICS is that the observer has to transform the energies from the observer's frame (L) to the electron's rest frame (E). Depending on the cross section regime, the characteristic energy of the ICS photons is equal to the seed photon energy boosted by a factor proportional to γ^2 or just γ, for the Thompson and Klein–Nishina regimes, respectively, where γ is the electron's Lorentz factor [198]. For low-energy photons scattered by relativistic electrons, the scattered photons, as seen in (L), are emitted in the same direction as their scatterers.

Inverse Compton emission occurs in astrophysical sites that are rich sources of relativistic electrons. According to certain models, which are reviewed later in this thesis, it is the mechanism chiefly responsible for the TeV γ rays from pulsars and their surrounding nebulae. Also, another very common source of IC emission are the microQSO jets, where background IR photons are upscattered by the ultra-relativistic electrons produced at the base of the jets. These objects are likely to be detected by the contemporary and future VHE experiments, provided their jets are aligned with the experiments' lines-of-sight.

2.7 Pair Production

2.7.1 On Matter

Pair production is the particle interaction mechanism that is chiefly responsible for the detection of VHE γ rays with our experiments. It is also the basis for models of high-energy emission in stellar environs (e.g. pulsar magnetospheres). VHE γ rays traversing the Earth's atmosphere are very likely to pair produce, provided they are exposed to the Coulomb field of a nucleus. This process is described as the complete replacement of the γ ray by an e^-–e^+ pair that carries the total energy of the photon. The theoretical photon energy threshold above which pair production takes place is 1.02 MeV, which is equal to the rest energy of the produced pair. In practice, it is likely that the parent photon will have higher energy than what is required for the process, in which case the energy excess is given to the pair in the form of kinetic energy, $2(\gamma - 1)m_ec^2$ [198]. This interaction is usually presented as

$$\gamma \xrightarrow{\epsilon_0} e^+ + e^- + E_{\rm kin} \tag{2.19}$$

It seems straightforward how the energy conservation is translated for pair production. However, the role of a nucleus in this interaction can be made clearer if we consider the closed system in which the high-energy photon is materialised into an e^-–e^+ pair. In this system, the total energy and momentum must be conserved. For simplicity, we consider here the case where the products propagate in the parent photon's direction:

- Momentum conservation

$$\frac{\hbar\omega}{c} = 2\gamma m_e v \tag{2.20}$$

- Energy conservation

$$\hbar\omega = 2\gamma m_e c^2 \tag{2.21}$$

where v is the velocity of the pair's particles, and ω is the angular frequency of the photon. It can be seen that the above equations cannot be both valid at the same time, because in such case they would violate either the energy or momentum conservation for the system (if we substitute Eq. 2.20 into Eq. 2.21 or vice versa, respectively). Hence, it becomes clear that an agent such as the electric field of a nucleus should be present in order to absorb some of the system's momentum. From equations 2.20 and 2.21, it can be deduced that the momentum excess needed to be absorbed by the third body is

$$\Delta J = 2\gamma m_e c(1-\beta) \tag{2.22}$$

where the relativistic $\beta < 1$. In practice, the angle between the photon's trajectory and that of the products decreases with increasing photon energy. At 30 MeV, the root-mean-square angle is roughly 4°, whereas in the VHE region it is $< 0.2°$. Therefore, for GeV γ-rays, the products of pair production follow a path almost parallel to the parent photon's direction.

It is also possible to have pair production in the presence of the ambient electric field of an electron, but the probability for this interaction is much lower than that for a nucleus. Moreover, the energy threshold for this interaction is 2.04 MeV [198]. A measure of the probability for pair production is given by the cross section of the interaction, and in general it depends on the parent photon's energy. Nevertheless, above the energy threshold that triggers Cherenkov emission in the atmosphere ($\gtrsim$ 50 MeV), it practically stays constant with energy and equal to

$$\sigma_{\rm pp} = \sigma_0 Z^2[(28/9)\ln(183/Z^{1/3}) - 2/27] \tag{2.23}$$

where Z is the atomic number of the nucleus, and $\sigma_0 = (1/137)e^4/(m_e c^2)^2$ [198].

Pair production becomes the dominant light–matter interaction above 30 MeV, and it is the photon-absorption mechanism exploited by space-borne γ-ray instruments like EGRET and *GLAST*, which have thresholds above this energy. From $\approx$ 100 keV and up to 30 MeV, Compton scattering is the most probable interaction (Fig. 2.7). COMPTEL is one example of instruments which utilise Compton scattering as part of their detection technique (see section 1.2).

Pair production is also the dominant interaction between the Earth's atmosphere and the VHE γ rays that are responsible for the Cherenkov air showers. The produced $e^- - e^+$ pairs create EM cascades by traversing the atmosphere at superluminal velocities, which produces low-energy Cherenkov light, and by annihilating with atmospheric electrons, which produces secondary γ rays that create subsequent EM cascades. We will not get into more detail at this point, but further discussion takes place in section 4.1.

2.7.2 On Low-Energy Photons

In addition to the, more common, pair production in the presence of matter, high-energy photons can also interact with other lower-energy photons, provided that the interaction's cross section permits it. This photon–photon interaction is usually represented as

$$\gamma + \gamma \xrightarrow{E_{\gamma\gamma}>2(m_ec^2)^2} e^+ + e^- \tag{2.24}$$

where $E_{\gamma\gamma}$ is the centre-of-momentum energy of the two-photon system, and m_ec^2 is the rest mass energy of the produced e^-/e^+. The probability for this interaction peaks just above its threshold, which is defined by the condition

$$E_\gamma = \frac{\left(m_ec^2\right)^2}{h\nu} \approx \frac{0.261}{h\nu}\ \mathrm{MeV} \tag{2.25}$$

where E_γ is the γ-ray photon's energy, and $h\nu$ is the energy of the low-energy photon. At this energy, the cross section is roughly 10^{-25} cm^2 [198].

Eq. 2.25 implies that there can be certain photon energy combinations which could lead to pair production. For example, such possibilities may include the interaction of 100-MeV γ rays with $\sim$ 5-keV X-ray photons, that between 1-TeV γ rays and near-infrared 0.5-eV photons, or some other combination. The interaction between IR and γ-ray photons is very common in the description of radiation mechanisms in pulsar magnetospheres where the high density of both γ and IR photons makes this process dominant. Further discussion on such processes can be found in section 3.1.3.

An important consequence of the photon–photon interaction for VHE γ-ray astronomy is that it limits the distance up to which γ rays can be observed without being attenuated by the background optical photons. These photons are abundant inside our galaxy, but, more importantly, they occupy the extragalactic space [1].

2.7.3 In Strong Magnetic Fields

In extreme astrophysical situations, where the magnetic fields are super-strong, yet another more exotic interaction can become important. The interaction of single photons with magnetic fields can result in pair production, provided that $B \gtrsim 10^9$ G and that the photons exceed the threshold energy, $E_\gamma \geq 2m_ec^2$. Such process is usually represented by

$$\gamma + \boldsymbol{B} \longrightarrow e^{+} + e^{-} \tag{2.26}$$

and imposes serious limitations on the propagation of high-energy γ rays through strong magnetic fields. A calculation of the attenuation that such photons undergo in a magnetic field has been done by Erber, and it shows that the mean free path for a photon in a magnetic field is inversely proportional to the magnetic component of the field which is perpendicular to the photon's direction. Consequently, photons crossing the magnetic field lines at a high pitch angle are more likely to pair produce than those travelling tangentially to the field lines [202].

Although this kind of interaction is theoretically possible for all photons with energies above the threshold, it can be shown that its occurrence becomes important only if $E_{\gamma}/(2m_{e}c^{2}) \times B_{\perp}/B_{\mathrm{crit}} \gtrsim 0.1$, where $B_{\mathrm{crit}} = 4.414 \times 10^{13}$ G is the critical magnetic field for which the gyroenergy ($\hbar\omega_{\mathrm{g}}$) of an electron is equal to its rest mass energy. Such powerful magnetic fields exist close to pulsar magnetic poles, a fact which gave rise to many theoretical models of pulsar high-energy emission (see section 3.1.3). However, as one moves away from the pulsar surface, the magnetic field strength can drop by as much as an order of magnitude, which makes magnetic pair production more improbable.

More recently, interest has turned towards even more exotic quantum-mechanical predictions with regards to light interactions in strong magnetic fields. Harding and Baring investigated the idea of magnetic photon-splitting in the magnetospheres of highly magnetised pulsars, like PSR B1509−58, which has the strongest surface magnetic field ($B \approx 3 \times 10^{13}$ G) amongst the γ-ray pulsars discovered so far. In short, this interaction can be represented as

$$\gamma \xrightarrow{B} \gamma\gamma \tag{2.27}$$

and is only important in magnetic fields in excess of $B \sim 10^{13}$ G. It predicts, as its name suggests, the splitting of a single high-energy photon into two lower-energy ones. It is also thought to be amongst the reasons why there have not been any discoveries of HE γ-ray pulsars with $B \gtrsim 10^{14}$ G [203].

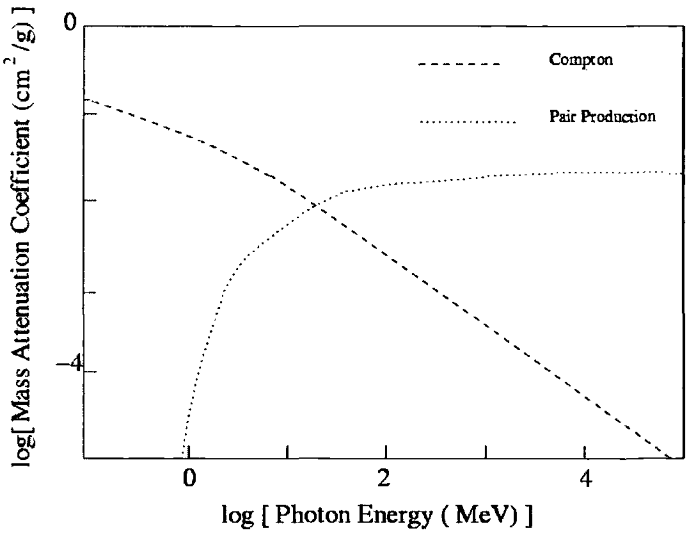


Figure 2.7: (from [1]) The mass attenuation coefficients — a quantity equivalent to the relative cross sections — for Compton scattering and pair production, in different energy ranges. Compton scattering (dashed line) is the dominant process up to 30 MeV and is the principle on which the Compton telescopes, like *INTEGRAL* and *CGRO*, operate. At higher energies, pair production is the most effective way of detecting γ-ray photons. In the VHE regime, the cross section reaches asymptotically the constant value σ_{pp}, discussed in the text.

Chapter 3

Pulsars: A Mystery Still Spinning in our Heads

3.1 Pulsars as Isolated Objects

3.1.1 Introduction

Our Sun is quite an average star. There are many others with properties, like stellar mass, magnetic field, temperature, etc., far greater and more impressive than the Sun's and many that are ranked below it. For a star about as massive as the Sun, the final stages before all its fuel source runs out involve the red giant phase, in which the star's radius expands up to 100-fold and the surface temperature drops by thousands of K, followed by the white dwarf phase, in which the star's core becomes degenerate under the high gravitational pressure.

However, there are stars a lot more massive than the Sun, which end their lives in a more spectacular fashion. For a star $\sim$ 15–30 times more massive than the Sun, the gravitational pressure is high enough to trigger not only the production of ^{12}C and ^{16}O, but also the fusion of ^{16}O to ^{32}S and, later, to ^{56}Fe, which has the highest binding energy per nucleon. At this point, there is no radiation pressure from burning to stop the collapse of the ^{56}Fe core, and as a result all nuclei are crushed together and the free electrons form a plasma state. Now, gravity's only opposing force is the degenerate pressure of the plasma electrons. For less massive stars, this is enough to bring the system to equilibrium, and the star ends its life as a white dwarf. In the case of massive stars, the degeneracy pressure is not enough to prevent the collapse, and unavoidably matter is further compressed. Under such tremendous pressures, the core's temperature rises to several billion K. Photons at such temperatures are so energetic as to be capable of splitting the ^{56}Fe and ^{4}He cores until only nucleons remain: this process is called photodisintegration. The photodisintegrated protons can now capture plasma electrons and form neutrons and neutrinos, i.e.

$$e^- + p^+ \longrightarrow n^0 + \nu_e \tag{3.1}$$

Without the electron degeneracy pressure, however, there is nothing to stop the core's free fall on itself, and so it is crushed down to an area of $\approx$ 25 km in diameter, within seconds. All the gravitational energy during the core-collapse ($\sim 10^{53}$ erg) is converted to internal energy of the core, while copious amounts of neutrinos are produced from photodisintegrations. After a certain density threshold, corresponding to $\sim 10^{11}$ g cm^{-3}, the core becomes opaque to these neutrinos, and as a consequence they get trapped within the core's boundaries. The mean free path of the trapped neutrinos is much shorter than the core radius and does not allow them to escape in time-scales equal to the system's dynamical time: equal to the duration of the core collapse. The core's spectacular collapse can only be halted, now, by the neutron degeneracy pressure. Although this is quite an efficient braking mechanism against gravity, the core bounces back in an effort to reach hydrostatic equilibrium [204]. This creates a shock wave which propagates towards the in-falling stellar matter, which did not have time to catch up with the core's collapse. During this propagation, the previously trapped neutrinos have now enough time to escape. The time, over which this happens, is equal to the diffusion time $\tau_{\rm diff} \approx 3$ s. During $\tau_{\rm diff}$, $\approx$ 99% of the energy that was stored in the core during the collapse is carried away by neutrinos. Approximately 1% is converted into kinetic energy of the outwardly propagating shock wave, 1% of which are cosmic rays. Finally, only 0.01% is transformed into optical emission, which is, however, enough to outshine the luminosity of the whole host galaxy. It is worth mentioning that the neutrino luminosity during $\tau_{\rm diff}$ is $\sim 3\times10^{19}L_\odot$, which is comparable to the optical luminosity of the entire visible universe [205]!

The collision of the in-falling matter with the expanding shock wave creates a spectacular supernova type II which expels matter into outer space at relativistic velocities. What is left is a star whose existence relies on the balance between gravity and neutron degeneracy pressure: i.e. a neutron star. Its mass can be somewhere between $\approx$ 1.4 and 2–3 $M_\odot$, where the lower limit corresponds to the Chandrasekhar limit for white dwarfs [206]. However, if a star in its end state is more massive than the above upper limit, the core simply collapses under gravity to an infinitely small area and forms a black hole.

Neutron stars are the most compact objects known to humans. Their size is that of a large city, and their densities exceed 10^{14} g cm^{-3} (the nuclear density). Hence, they can be conceived as giant nuclei composed of neutrons which are held together, not by the strong force, but by gravity. In the words of Frank Shu, "A sugar cube of neutron star-stuff on Earth would weigh as much as all of humanity! This illustrates again how much of humanity is empty space." [207].

However, these places of extreme densities and gravitational fields had not always been that way. The primordial material that formed those massive stars was lumped together, under gravity, from a vast region in which densities are lower than in any of the artificial vacuum humans can create. Although seemingly unconnected, there are properties that the primordial molecular cloud inherited to the neutron stars observed today. In time-scales of Gy, gravitational attraction overpowered the stochastic molecular motions of the gas and forced the primordial gas to contract via the gravitational paths defined by Kepler's laws. Ultimately, this led to the spinning gaseous spheres we observe as stars (see Fig. 3.1). Furthermore, "Biermann's battery" effect, plus other "dynamo" mechanisms, acted on the stellar interior and strengthened the initial magnetic field to the observed stellar values [208]. For giant OB stars, which are the likely progenitors of neutron-star formation, the surface magnetic fields can reach $\sim 1,500$ G. Also, it is believed that the dynamo mechanism responsible for this value is of different origin than that for solar-type stars [209].

Consider a main sequence star like the Sun, with an average period of rotation — from the equator to the poles — $P_\odot \approx 2.5 \times 10^6$ s, and an average surface magnetic field $B_\odot \sim 1$ G. Assuming conservation of angular momentum and magnetic flux for a star destined to become a supernova type II, we can qualitatively derive the following values for a neutron star:

- Angular momentum conservation

$$MR^2\Omega = const. \Rightarrow P_{\rm ns} = P_\odot \left(\frac{R_{\rm ns}}{R_\odot}\right)^2 \sim 5 \times 10^{-4}\ {\rm s} \tag{3.2}$$

- Magnetic flux conservation

$$BR^2 = const. \Rightarrow B_\mathrm{s} = B_\odot \left(\frac{R_\odot}{R_\mathrm{ns}}\right)^2 \sim 5 \times 10^9 \text{ G} \tag{3.3}$$

where $R_\odot = 7 \times 10^8$ m and $\omega = 2\pi/P$ are the solar radius and angular frequency of rotation, respectively. The observed typical values for neutron-star rotation periods start from $\sim 10^{-3}$ s. This simplified calculation, however, shows that in theory such fast rotations are possible. In addition, the typical magnetic fields of such stars can easily reach 10^{12} G. As mentioned earlier, the 10^3-factor difference could be accounted for by the stronger magnetic fields present in the more massive OB stars.

In these first few paragraphs, the formation and basic properties of a neutron star were introduced. These stars are not just a theorist's imagination, although their existence was proposed by Zwicky and Baade in 1934 as a purely theoretical hypothesis [210]. This hypothesis was verified in 1967, when neutron stars were unexpectedly observed for the first time with the Cambridge, 4-acre radio telescope. At the time of their discovery, Antony Hewish and his research student, Jocelyn Bell, were using the telescope for a totally different area of radio astronomy, which aimed at the detection of compact radio sources using the solar wind's influence on their radio signals. What Hewish and Bell detected was the pulsed radio emission from CP 1919 [211], which is a highly magnetised, rapidly rotating neutron star (see Fig. 3.2). Such a star is called a pulsar, and it is the aim of the following paragraphs to describe, in more detail, its physical properties as well as the theoretical predictions of pulsed γ-ray emission from these objects.

3.1.2 Geometry and Physical Properties

The Interior

There exists a number of models that try to describe the interior of a neutron star. The basis of all models is the equation of state (EOS) that describes the dependence of pressure, p, on matter density, ρ, inside a neutron star. The ones that are characterised by 'soft' EOSs assume an attractive force, on average, between the nucleons, whereas models that are based on 'stiff' equations of state allow a repulsive potential at the inner parts of a neutron star. More theoretical support exists for the latter, which lead to neutron stars with larger radii [184]. Regardless of the differentiation between models, a general consensus has been reached, which has led

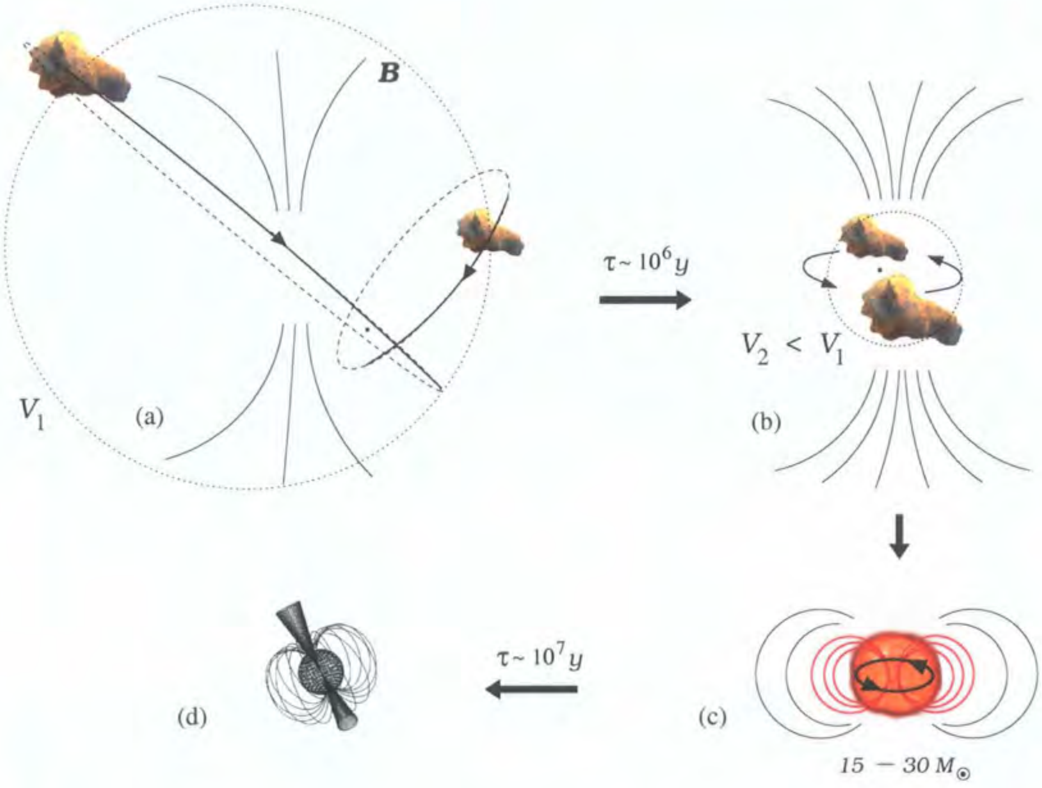


Figure 3.1: There are regions in space, where interstellar matter forms distributions of molecular clouds that are held together under gravitational attraction. These distributions can maintain their volume (V_1), if gravity balances the stochastic motions of the molecular gas, which force the system to expand. However, if instabilities start to appear on a large enough scale, (a) the gaseous masses will be forced to follow highly eccentric, Keplerian orbits towards the system's centre of gravity [212]. (b) Further contraction to a smaller volume (V_2) triggers the formation of rotating spherical distributions of gas. Meanwhile, the interstellar magnetic fields are enhanced via the "Biermann's battery" effect (described in [208]). Once contraction has adiabatically compressed the gas to the level where nuclear reactions can start, star formation begins. (c) Beyond this stage, it is believed that stellar magnetic fields are strengthened even further, via an internal "dynamo" mechanism. Especially for massive OB stars, this leads to surface magnetic fields that can be $\sim 1,500$ G. The magnetic field origin inside these stars is thought to be different to that of solar-like stars [209]. (d) After $\sim 10^7$ y, a massive $\sim$ 15–30-$M_\odot$ star exits the main sequence and via a spectacular, type II supernova explosion is likely to become a rapidly spinning, highly magnetised neutron star, also known as a pulsar.

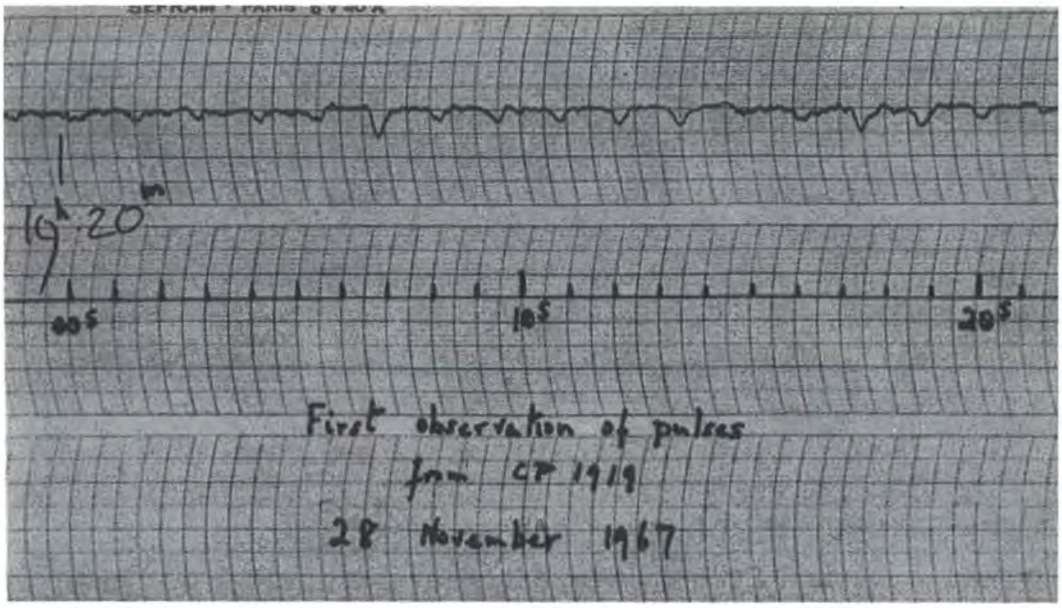


Figure 3.2: (from [184]) Pulsations from CP 1919 (PSR B1919+21), detected with the Cambridge, 4-acre radio telescope by Antony Hewish and Jocelyn Bell. The pulses on this graph are shown as downward deflections of the trace.

to the following description of the neutron star's interior:

> It was mentioned earlier that the neutron star remnant is the result of photodisintegration of iron nuclei and the subsequent electron-capture by protons, towards the formation of neutrons. So, in many ways, a neutron star can be considered as a giant nucleus consisting of $\sim 10^{60}$ neutrons [98]. In reality, that describes only the inner layers of a neutron star's composition. As we move from the central layers to the surface, the total force per surface unit — the result from adding the gravitational pressure to the, opposite, degeneracy pressure — decreases, and therefore p decreases. That allows for different states of matter to exist throughout the stellar body. A simple schematic of the distribution of matter is shown in Fig. 3.3. It can be seen that a large part of the stellar interior is thought to be superfluid. Such fluid has no viscosity and is able to rotate independently from the rest of the neutron star. However, because the transition between the crust and the core is gradual, it is thought that there exists a significant amount of electrons and protons on either side of the border that lies between them. This important property of the neutron

star interior has been connected with an unusual feature of a pulsar's timing behaviour, the *glitches*.

The glitches are irregularities in the otherwise stable timing behaviour of a pulsar. They are described as sudden spin-ups of a pulsar's rotation, followed by an exponential recovery. The theoretical background for the observed glitches has been laid down by Shapiro and Teukolsky, who provided an analytical expression that describes the timing behaviour of a pulsar during, and shortly after, a glitch. Their solution, which is called the glitch function, gives the pulsar's angular frequency, ω, as it changes during a glitch: i.e.

$$\omega(t) = \omega_0(t) + \Delta\omega_0 \left[Q \exp(-t/\tau) + 1 - Q\right] \tag{3.4}$$

where $\omega_0(t)$ is the angular frequency function that characterises the pulsar's *normal* timing behaviour: i.e. without the glitch. $\Delta\omega$ is the angular frequency amount, by which the pulsar was spun up during the glitch. Q is the so-called healing parameter, which represents the degree to which the angular frequency is recovered to the value that the pulsar would have normally had without the glitch: i.e. the one that results from the extrapolation of ω_0. Finally, τ is the relaxation time of the exponential recovery and is the one parameter which has been connected with the physical properties of the neutron star interior.

The TeraGauss magnetic fields that dominate a pulsar's surface and interior are believed to be responsible for coupling the superfluid's rotation, via its charged component, with the crust's rotation. However, this coupling becomes weaker on the border between the crust and the fluid core, and as a result the rotation becomes differential. It is believed that this motion can affect pulsar rotation and account for the observed sudden spin-ups (glitches). The time-scale for frictional dissipation between these two components and the amount of time during which they exchange angular momentum has been assigned to τ. Hence, according to this two-component model, we have a directly measurable quantity that could provide an insight into pulsar structure [98],[184].

The Magnetosphere

A pulsar's *magnetosphere* occupies the space beyond the stellar surface, where magnetic fields dominate the underlying physics. It is commonly accepted now that the immediate environment of a pulsar is filled with electron and proton plasma.

However, when the idea of a neutron star was first conceived, it was thought that the strong gravity would not allow for any particle to escape the stellar surface, which meant that pulsar atmospheres must be very thin [213]. Nevertheless, in such a case it turns out that the magnetic field on the surface is such that the induced electrostatic forces exerted on the surface particles would be enormous. As will be seen in the next paragraph, the magnitude of these forces is roughly 10^{10} times stronger than gravity and hence capable of pulling particles off the surface of the star and accelerating them away from it. This mainly occurs above the *polar caps*, which are highly conductive regions around the magnetic poles. There, the magnetic field lines are perpendicular to the pulsar surface and allow a virtually frictionless propagation of charge along them, but *not* across them. According to one of the popular theoretical models, polar caps are the generators of high-energy emission from pulsars.

The magnetic field lines together with the charged particles of the magnetosphere are able to follow the star's rapid rotation, only up to a certain height above the surface. There is a point, however, beyond which the radial velocity of the particles becomes equal to c and therefore matter cannot follow the rotation. The boundary at which this happens is a cylindrical surface called the *light cylinder*, and its radius is defined by c/Ω, where Ω is the pulsar's angular frequency (see Fig. 3.5). The entire magnetosphere is then contained inside the volume defined by the light cylinder, from where all the observed pulsar emission is believed to originate. The field lines of the magnetosphere can be divided into "closed" and "open". Open magnetic field lines emanate from high magnetic latitudes inside the polar cap regions and accelerate particles beyond the light cylinder, where these particles are disconnected from the magnetosphere. Therefore, these field lines — defined by the particle trajectories along them — do not close inside the magnetosphere, but extend beyond it. Also, open magnetic field lines give rise to pulsar wind — i.e. the flow of particles *towards* and *from* the pulsar magnetosphere — which connects it with its surrounding space. On the other hand, closed field lines are those completely contained inside the light cylinder and are able to follow the pulsar's rotation up to the light cylinder's edge. Hence, there exists a last, closed magnetic field line which is tangential to the light cylinder. The region on the pulsar's surface where these lines originate from defines the border of the polar caps, and it is a circle of radius

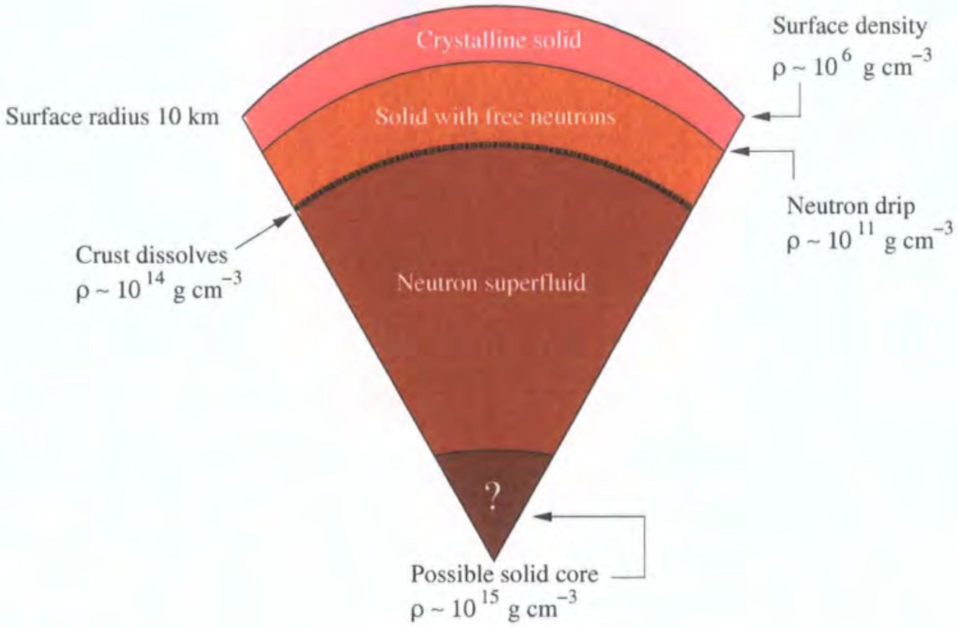


Figure 3.3: (from [184]) Model of neutron star structure. In the outermost layer, the density ($\sim 10^6$ g cm^{-3}) is not high enough for the ^{56}Fe nuclei to photodisintegrate, and so the surface of the neutron star is a strong, crystalline ^{56}Fe crust. As one goes deeper into the star, densities rise to $\sim 10^{11}$ g cm^{-3}, and now the nuclei start to split, which results in neutron formation ("neutron drip" point). In the star's core, the densities reach $\sim 10^{14}$ g cm^{-3}, and matter is so highly compressed that no nuclei can exist and neutron plasma is formed. This degenerate state of matter is a super-conducting fluid with no viscosity. It is thought, however, that in the innermost parts of the star, the neutron superfluid can be further compressed to nuclear densities of $\sim 10^{15}$ g cm^{-3}.

$$R_{\mathrm{pc}} \approx \sqrt{\frac{R_{\mathrm{ns}}^3}{R_c}} \quad \text{if } R_{\mathrm{ns}} \ll R_c \tag{3.5}$$

where R_{ns} is the neutron star's radius and $R_c = c/\Omega$, the light cylinder's radius [214].

The Charge-Separated Magnetosphere

The first treatment of a pulsar's electrodynamics was given by Goldreich and Julian, where it was assumed that the neutron-star interior is a perfect conductor, and the surrounding magnetosphere is the extension of the star's internal magnetic field

[215]. Goldreich and Julian investigated the simple case of an *aligned rotator* — a pulsar whose rotation and magnetic axes are aligned — and they concluded that pulsars are surrounded by charged plasma. Otherwise, as they showed, the boundary conditions for the electric and magnetic fields on the stellar surface would lead to an enormous induced electric field, which is capable of stripping the neutron star of its surface charge. Such conditions are obviously unstable, and thus the vicinity of the neutron star must be charged as well.

This statement has some interesting implications that helped the above authors to construct an image of the pulsar magnetosphere. They considered the internal, dipolar magnetic field of the star, which in the case of a pulsar rotates with angular velocity, Ω, and stipulated that since the neutron star is a perfect conductor, the induced electric forces, $\boldsymbol{E}e$, caused by the rotating magnetic-dipole field should cancel out the magnetostatic forces, $e(\boldsymbol{\Omega} \times \boldsymbol{r}) \times \boldsymbol{B}$. So that everywhere inside the neutron star

$$\boldsymbol{E} + \frac{\boldsymbol{\Omega} \times \boldsymbol{r}}{c} \times \boldsymbol{B} = \boldsymbol{0} \tag{3.6}$$

where $\boldsymbol{B}$ is the magnetic field at distance r from the centre of the star, in Gauss, and $\boldsymbol{E}$ is the induced electric field.

Moreover, based on the boundary conditions for a conductor, they showed that if one assumes that the neutron star is surrounded by vacuum, the induced electric field on protons of the polar cap region overpowers gravity by a factor of $5 \times 10^8 B_{12} R_6^3/(PM)$. Here, $B_{12} = B/(10^{12}\ \mathrm{G})$, $R_6 = R/(10^6\ \mathrm{cm})$, P is the rotation period in seconds, and M is the stellar mass in number of solar masses. For a 10-ms pulsar of 1 $M_\odot$, with a typical surface magnetic field $\sim 10^{12}$ G and a 10-km radius, this factor yields 5×10^{10}. Therefore, as they concluded, the magnetosphere is charged, and they expanded this statement by saying that it is also the extension of the stellar surface and, hence, also a conductor. In other words, Eq. 3.6 should hold for both the interior and the magnetosphere of the neutron star. Such a condition leaves only the option of a co-rotating magnetosphere, where everywhere apart from special places, as we shall see, the Lorentz invariant, $\boldsymbol{E} \cdot \boldsymbol{B}$, vanishes. This implies that there is no net induced electric force that acts upon the charged particles which co-rotate with the magnetosphere (Fig. 3.4). Then, it follows from Eq. 3.6 that the charge density close to the pulsar surface, where $\Omega r \ll c$, is given by the Godreich–Julian density,

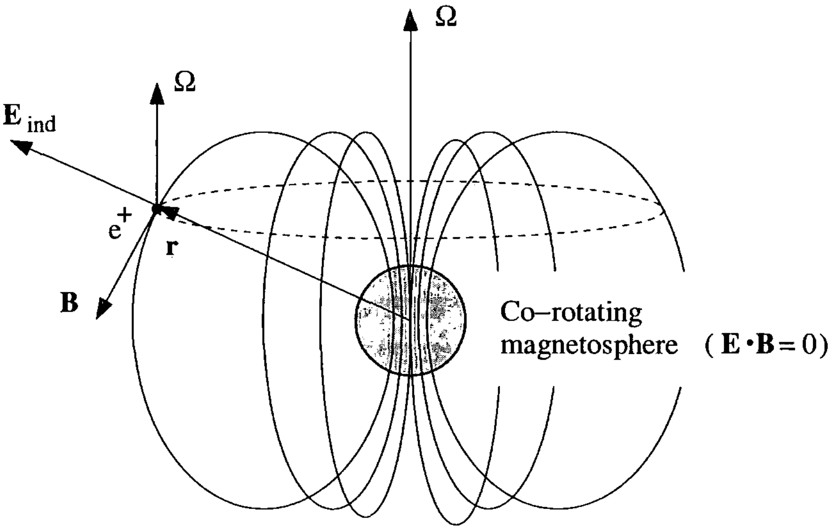


Figure 3.4: Electromagnetic forces in the co-rotating magnetosphere of an aligned rotator. Assuming that the pulsar magnetosphere is a perfect conductor, the magnetostatic forces must cancel out the induced electric field and force the charged magnetospheric plasma to co-rotate with the star.

$$\rho_{\mathrm{GJ}} = \frac{\nabla \cdot \boldsymbol{E}}{4\pi} \approx -\frac{\boldsymbol{\Omega} \cdot \boldsymbol{B}}{2\pi c} \tag{3.7}$$

In other words, the magnetosphere is charge-separated, with the borderlines between positive and negative regions being the surfaces defined by $\boldsymbol{\Omega} \cdot \boldsymbol{B} = 0$. Since there can be no charge flow through these surfaces, which makes them neutral, they have been appropriately named *null charge* surfaces (see Fig. 3.5).

In Julian and Goldreich's case of an aligned rotator, the electric potential is highest at the equator and decreases towards the poles. This means that the open field lines around the magnetic poles are streaming out protons, whereas at lower latitudes and, of course, inside the polar cap zones, the pulsar wind is mainly electrons. Despite the fact that this model is based on numerous assumptions, it gives a clear image of the pulsar geometry and is considered part of the standard model on which most of its successors are based. However, it is unable to provide an explanation for the observed pulses from these objects, a feat which other more modern models are trying to achieve. An overview of these models is given in the following paragraphs.

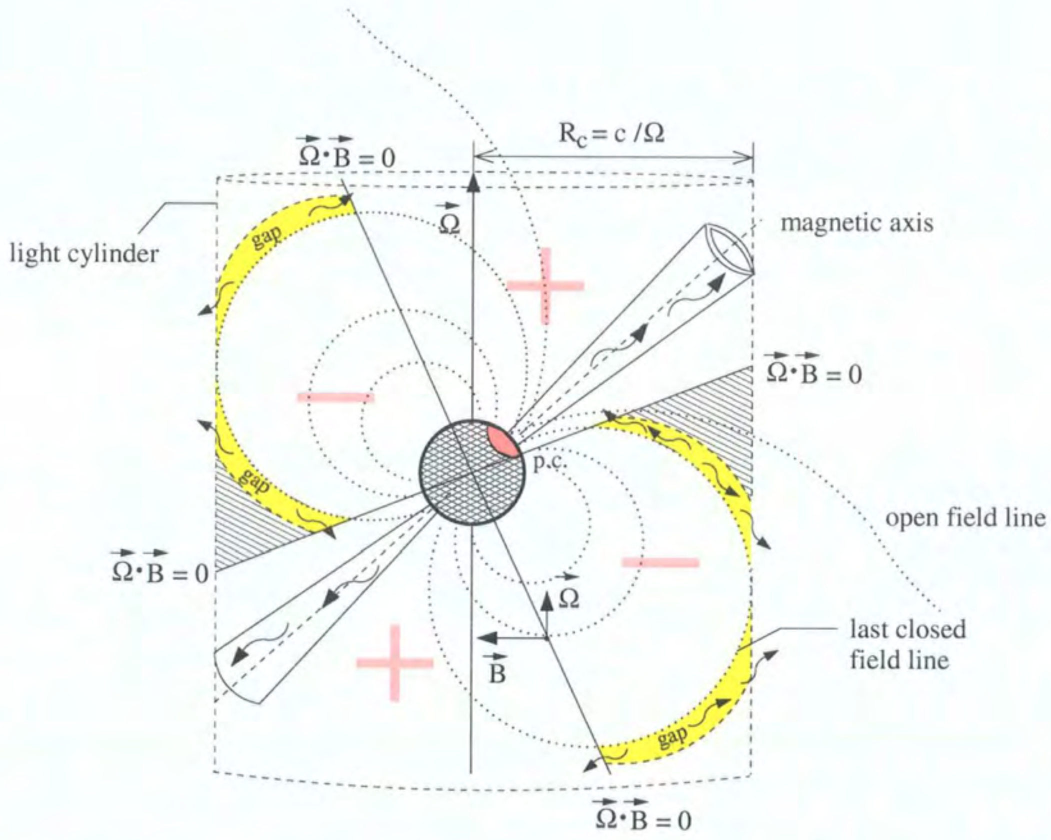


Figure 3.5: The geometry of the charged-separated magnetosphere of pulsars, and the regions where the radiation is generated according to the two established theoretical models: the Polar Cap and the Outer Gap.

3.1.3 Models of Pulsar High-Energy Emission

Established Views

The scientific interest on pulsars has been high since their discovery. In the past 38 years, there has been substantial theoretical and observational progress on these objects. However, the validity of theoretical models that are trying to describe the pulsar emission mechanisms is still a matter of further clarification through observations, especially in the VHE regime. Nevertheless, there are certain truths that we *do* know about pulsars, with regards to which a consensus amongst astronomers has been reached.

Soon after the first radio observations of pulsars, it was found that, in the pulsar's reference frame, the arrival times of radio pulses were characterised by a monotonically increasing delay. The delay was a tiny fraction of the pulsar frequency, but not negligible [216]. Hence, this meant that pulsars are slowing down. Theories that tried to explain the reasons behind these irregularities followed shortly afterwards [217],[215],[218]. Regardless of the details of each theory — which will not be mentioned here — the fact that pulsars slow down, together with their emission in a broad range of frequencies, has led us to the conclusion that part of their kinetic energy is converted into radiation. This conversion occurs via emission mechanisms that are yet to be clarified, and the effort to explain the observed pulse shapes and energy distributions is ongoing. Neutron stars whose emission is largely dependent on the conversion of their kinetic energy to observed radiation are called *rotation-powered* pulsars.

There are many interesting properties one can derive for rotation-powered pulsars. Their bolometric luminosity (power output) can be directly related to their spin-down which is measured from observations. Furthermore, under the assumption that pulsars are rotating magnetic dipoles in vacuum, one can give estimates of their surface magnetic fields and ages. Here, as an example, we will use the Crab pulsar, whose properties are presented in Table 3.1.

The kinetic energy of a solid, spherical neutron star that is rotating with frequency $f(t)$, where $\dot{f} < 0$, is

$$E_{\rm rot} = \frac{1}{2} I \Omega^2 = \frac{4\pi^2}{5} M_{\rm ns} R_{\rm ns}^2 f^2 \tag{3.8}$$

Since f decreases with time, the pulsar loses energy at rate

Surface magnetic field (B_s)	4×10^{12} G
Frequency of rotation (f)	30 Hz
First derivative of the frequency ($\dot{f}$)	-3.7×10^{-10} Hz s^{-1}
Neutron star radius (R_{ns})	10^6 cm
Neutron star mass (M_{ns})	1.4 $M_\odot$
Moment of inertia (I)	10^{45} g cm^2
Induced electric field (E)	6×10^{10} V cm^{-1}

Table 3.1: (from [1]) The basic properties of the Crab pulsar.

$$\dot{E}_{\rm rot} = \frac{8\pi^2}{5} M_{\rm ns} R_{\rm ns}^2 f \dot{f} \tag{3.9}$$

where $M_{\rm ns}$ and $R_{\rm ns}$ are the neutron star mass and radius, respectively.

Based on Ostriker and Gunn's assumption that the energy output from a rotation-powered pulsar is due to its magnetic dipole radiation, one can give estimates of the average surface magnetic field of a pulsar, $\langle B_{\rm s} \rangle$, by equating the spin-down power, $\dot{E}_{\rm rot}$, with the magnetic dipole radiation, $\dot{E}_{\rm brak}$; the latter is also called *braking radiation*[1] [218]. This radiation is of the order

$$\dot{E}_{\rm brak} \propto -m_\perp^2 f^4 \sim -\langle B_{\rm s} \rangle^2 R_{\rm ns}^6 f^4 \tag{3.10}$$

where $m_\perp$ is the magnetic dipole moment for an orthogonal rotator: a pulsar whose axis of rotation is perpendicular to the magnetic dipole's.

Setting equal equations 3.9 and 3.10 leads to

$$\langle B_{\rm s} \rangle = \sqrt{\frac{3c^3 I}{8\pi^2 R_{\rm ns}^6}} \cdot (P\dot{P})^{1/2} \tag{3.11}$$

For a typical pulsar, the proportionality constant before $(P\dot{P})^{1/2}$ is equal to 3.2×10^{19} G s$^{-1/2}$, if cgs units are used (see Table 3.1).

Finally, an estimate for the pulsar age can be given if one assumes that when the catastrophic supernova event took place the newly born neutron star was rotating at a frequency f_0, and immediately afterwards the magnetic dipole radiation — or, indeed, other mechanisms through which the pulsar loses kinetic energy — resulted in a gradual decrease of the frequency to the value we observe today. There is

---

[1] Not to be confused with the German term Bremsstrahlung (see section 2.3), which also translates as *braking radiation*.

experimental evidence that this gradual decrease is not as straightforward as Eq. 3.11 implies for a constant magnetic field [197],[219]. Part of this evidence comes from the famous plot that relates $\dot{P}$ with P, for various observed pulsars (see Fig. 3.6). Nevertheless, it is logical to assume that pulsars slow down faster when their energy content is larger, which is when their *frequency* is higher. In general, it is typical to assume a power law for $f(t)$, but not necessarily the one implied in Ostriker and Gunn's model, as that would immediately show a clear trend in the P–$\dot{P}$ diagram.

Starting from a generalised power law, we can conclude the following:

$$\dot{f} \propto -f^n \Rightarrow f^{-n}\dot{f} = -k \Rightarrow f^{-n}df = -kdt \tag{3.12}$$

where $k > 0$ is a proportionality constant, and $n > 1$ is the so-called braking index. In the case of Ostriker and Gunn's pure–dipole emission model, it is implied that $n = 3$ — since from Eq. 3.11 we have $P\dot{P} = \dot{f}/f^3 = k$ — and therefore $\dot{f} \propto f^3$.

Integration from time $t = 0$, when the frequency was f_0, to time t, when the frequency is f, gives

$$\frac{1}{n-1}\left(f^{1-n} - f_0^{1-n}\right) = kt = -f^{-n}\dot{f}t \tag{3.13}$$

After rearranging the terms, we get the following expression for the pulsar age:

$$t = -\frac{1}{n-1}\frac{f}{\dot{f}}\left[1 - \left(\frac{f}{f_0}\right)^{n-1}\right] \tag{3.14}$$

Given enough time for a pulsar to slow down substantially (i.e. $f \ll f_0$), we can approximate this expression to the more useful

$$\tau = \frac{1}{1-n}\frac{f}{\dot{f}} \tag{3.15}$$

This expression gives a rough estimate of the pulsar age, which is known as characteristic age, given the values of f and $\dot{f}$, and the braking index, n. The frequency and its first derivative can be measured for most pulsars, but the value of n is known for only a few, young ones. The difficulty in measuring n lies in its dependence on $\ddot{f}$. The conventional definition of the braking index is derived from Eq. 3.12 with differentiation. So, from that equation, by taking the logarithms of both sides and differentiating, we arrive at

$$-\dot{f} = kf^n \Rightarrow \log(-\dot{f}) = \log k + n \log f \Rightarrow \frac{\ddot{f}}{\dot{f}} = n\frac{\dot{f}}{f} \Rightarrow n = \frac{f\ddot{f}}{\dot{f}^2} \tag{3.16}$$

So far, there are only four pulsars whose braking indices have been measured. Two of them are the EGRET γ-ray pulsars, Crab and Vela. The other two are PSR B1509−58, which has been detected with COMPTEL, and the Crab-like PSR B0540−69, which was discovered in soft X-rays with the *Einstein* satellite. For all of them, the braking indices are below the value which corresponds to dipolar emission and range from 1.4 to 2.8 [220],[221],[222],[223]. However, a consensus of accepting magnetic dipolar emission (i.e. $\langle B_\mathrm{s} \rangle = \text{const.} \Leftrightarrow n = 3$) prior to defining characteristic age has been reached, which results in $\tau = -(1/2)f/\dot{f} = (1/2)P/\dot{P}$. We can now apply this equation to the case of the Crab pulsar — one of the few pulsars whose birth has been documented — and compare the result with this pulsar's actual age.

On 4 July 1054, Chinese astronomers recorded the appearance of an abnormally bright star in the constellation of Taurus, which shone — at its brightest — 4 times brighter than Venus. According to the records, the star continued to shine day and night for 23 days [224]. We know now that this "guest star", as was characterised by its first observers, was at the location where one can find the Crab nebula: a supernova remnant that harbours the Crab pulsar. This rare historical information can help us check the validity of the magnetic dipolar emission model, for this case.

Given the values of Table 3.1 for the Crab pulsar, Eq. 3.15 gives an estimate of $\sim$ 1300 y. This value is significantly different ($\approx 35\%$) from the documented 950 y, but there were many assumptions prior to the calculation, which may not be true. Perhaps the magnetic field of the pulsar has degraded with time [225], or there exist other mechanisms which removed angular momentum from the pulsar (e.g. gravitational waves). As a result, the assumed power law is different. Gravitational emission cannot be checked experimentally yet, but if it does occur, this age discrepancy can be removed by adjusting the parameters of the theory [218].

Let us consider now the magnetic field of the Crab pulsar. According to Eq. 3.11, the average surface magnetic field is $\langle B_\mathrm{s} \rangle = 3.7 \times 10^{12}$ G. A comparison of this value with the one from Eq. 3.3 reveals that the pulsar energy output is largely based on its progenitor star's properties. This energy output for the Crab pulsar is $\dot{E}_\mathrm{rot} = \dot{E}_\mathrm{brak} = 5 \times 10^{38}$ ergs s^{-1} and matches the order of magnitude of the bolometric luminosity observed from the entire Crab nebula. In other words, the

Crab pulsar, a 10-km-wide neutron star, can be regarded as the power source of a vast, 10-ly-wide region (see Fig. 3.7). Exactly where this energy source is located around the neutron star is a matter of large debate, and there are two dominating theories, which place it either close to the light cylinder or in the acceleration zones over the polar caps.

Observability and γ-ray Efficiency

The expression for the spin-down luminosity in Eq. 3.9 can help us define another useful quantity: the pulsar *observability*. The observability of a pulsar that lies at a certain distance (say d) is equal to its spin-down energy flux that would be observed if the emission were uniformly distributed over the whole sky. Hence, the observability is defined as

$$L_{\text{obs}} = \frac{\dot{E}_{\text{rot}}}{4\pi d^2} \tag{3.17}$$

and it is a measure of how bright a pulsar is expected to appear, based on its spin-down luminosity $\dot{E}_{\text{rot}}$.

A remarkable fact is that 6 out of 7 high-energy pulsars, observed with EGRET and COMPTEL, have a top rank according to observability amongst the discovered pulsar sample (see Table 3.3). PSR B1055−52 is ranked 29th, which can still be considered high in a sample of more than 1,300 pulsars [227]. However, this raises an important question: why have the rest of the radio pulsars, ranked above PSR B1055−52, not been observed in γ rays? It seems that the answer is not simple, since it could involve the relative geometry between the observer's line-of-sight and the direction of γ-ray emission, which could potentially conceal the pulsar high-energy radiation. This subject is mainly a concern for the theoretical models of pulsar emission. Some of these models are discussed in the following paragraphs. Nevertheless, the issue can be dealt with by focusing only on pulsar energetics.

Eq. 3.17 does not consider the type of emission which a pulsar's rotational energy is converted to. It can be argued, therefore, that one of the reasons behind the detection of only 7 high-energy pulsars, or equivalently the non-detection of radio pulsars with high observability, is the ability of the latter to convert kinetic energy into γ rays. This ability is simply expressed with the ratio between pulsar γ-ray luminosity, L_γ, and spin-down luminosity, $\dot{E}_{\text{rot}}$, which is called γ-ray *conversion efficiency*. Hence, the γ-ray conversion efficiency is defined as

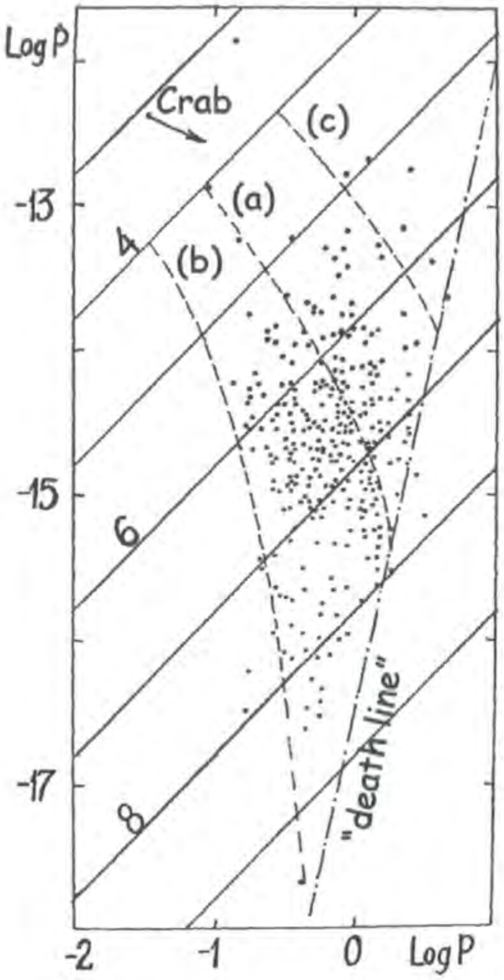


Figure 3.6: (from [219]) Several observations of pulsars show a wide spread of values on the P–$\dot{P}$ diagram. If there is a law that describes how pulsars slow down with time, it does not appear to be straightforward. It is therefore possible that different mechanisms act together towards the loss of kinetic energy from a pulsar. The dashed lines (a), (b) and (c) in the above P–$\dot{P}$ plot show three evolutionary tracks: (a) is the track that most pulsars follow; depending on their mass, magnetic field and rotational velocity at birth, pulsars could follow steeper (b) or more gradual evolutionary tracks (c) towards the "death line" (dashed–dotted line), beyond which most cease to emit radiation. The current evolutionary direction of the Crab pulsar is indicated with the arrow. The diagonally distributed, numbered solid lines in this plot are characteristic-age contours, given by $\log(\tau/y) = \log[P/(2\dot{P})]$.

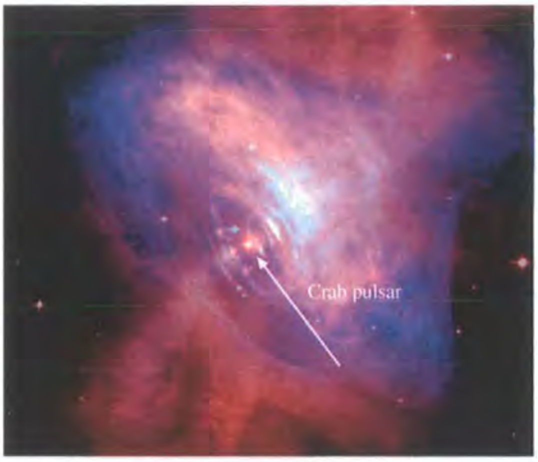


Figure 3.7: (from [226]) The Crab nebula and pulsar as seen in optical and X-rays with the *Hubble Space Telescope* and the *Chandra* X-ray satellite, respectively.

$$\epsilon_\gamma = \frac{L_\gamma}{\dot{E}_{\rm rot}} \tag{3.18}$$

This definition of the conversion efficiency can be directly connected to the experimental results from EGRET observations, if one assumes a beaming factor, $f_\gamma = \Omega/(4\pi)$, which is equal to the fraction of the sky that is subtended by a pulsar's beam. Since f_γ is something that cannot be measured, it is usually assumed equal to 1 sr. Hence, a more practical definition of ϵ_γ is

$$\epsilon_\gamma = 4\pi d^2 f_\gamma \frac{F_\gamma}{\dot{E}_{\rm rot}} \langle E \rangle \tag{3.19}$$

where F_γ is the measured integral γ-ray flux over an energy range $[E_1, E_2]$, and $\langle E \rangle$ is the mean γ-ray energy over the observed range.

γ-ray conversion efficiencies have been calculated for all EGRET pulsars (see Table 3.3). Considering the low observability of PSR B1055$-$52, which is ranked below 22 other radio pulsars not detected by EGRET, it is bourne in mind that a relatively weak overall emitter could still appear as a bright γ-ray pulsar. In the case of PSR B1055$-$52, the conversion efficiency is $\sim 15\%$ (the highest amongst EGRET pulsars), which compensates for its low ranking.

The age-spread of EGRET pulsars over ≈ 3 orders of magnitude was the motivation for correlation studies between pulsar γ-ray efficiency and age. Despite the limited available sample, there is a clear trend that reveals an increased γ-ray efficiency for older pulsars. Zhang, Cheng, and Rudak and Dyks fitted the efficiencies of all 6 EGRET pulsars (see Fig. 3.8) and found that the phenomenological function that represents γ-ray efficiency evolution with time has the linear form [228],[229]

$$\epsilon_\gamma \approx 1.3 \times 10^{-7} \tau \tag{3.20}$$

Although this effort looks promising, the statistical uncertainties are large. A more complete picture will require a larger sample, which is expected, however, to arise from EGRET's successors (e.g. *GLAST*) but also from ground-based telescopes (e.g. MAGIC), since their energy thresholds and sensitivities are much improved with respect to EGRET.

Polar Cap Models

Although the basic principles of a pulsar's magnetospheric structure had been laid down, the explanation and prediction of the observed pulses and spectra required a more rigorous treatment of the underlying physics. Early models assumed that the polar-cap zones were the source of electromagnetic radiation from pulsars (see e.g. [230]). The driving force behind this radiation was assumed to be the electrostatic potential drop between the pulsar surface and the regions above. The main mechanisms involved in Polar Cap models are curvature and synchrotron radiation mainly from electrons and positrons, as well as from protons in some cases. Some of the main contributors towards this class of models were P. A. Sturrock, M. A. Ruderman, P. G. Sutherland, A. K. Harding and J. K. Daugherty [230],[231],[232]. An overview of the different approaches is provided here for the reader.

In Sturrock's model, there is a region above the polar caps, where the Lorentz invariant, $\boldsymbol{E} \cdot \boldsymbol{B}$, does not vanish. Thus, there is a net induced electric field, $\boldsymbol{E}(h)$, which is strong enough to pull ions out of the surface and accelerate them across a height difference, $h \approx R_{\rm pc}$, above the pulsar surface. Sturrock considered that as one moves radially away from the polar cap the magnetic field lines change gradually from radial to almost transverse to the electric field. However, he assumed for simplicity that up to the above critical height the electric field is parallel to the magnetic field lines, which gives rise to effective particle acceleration. Beyond that point, and everywhere else in the magnetosphere, Sturrock assumed that the magnetic field is normal to the electric field (i.e. $\boldsymbol{E} \cdot \boldsymbol{B} = 0$), and the charged particles follow the magnetic field lines closely — and, thus, have negligible transverse momentum —

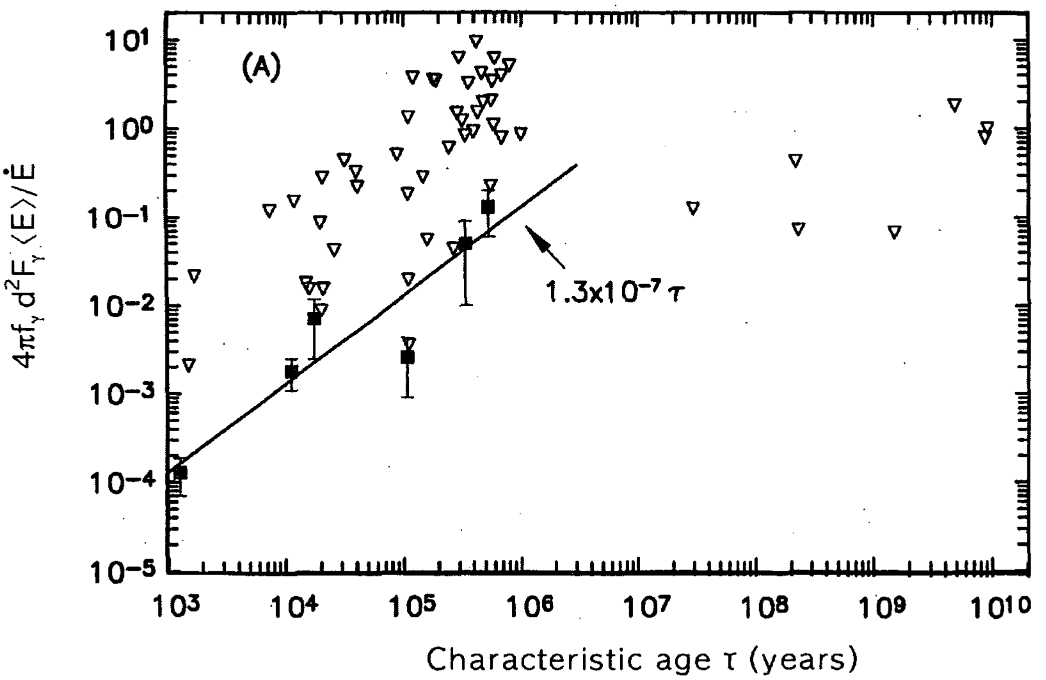


Figure 3.8: (from [228]) This plot shows that there seems to be a phenomenological function for EGRET pulsars, which relates γ-ray efficiency with characteristic age. However, the limited statistics as well as the complex nature of γ-ray production make this assumption premature. Zhang and Cheng selected 57 pulsars compliant with the Outer Gap model, only 6 of which have been clearly detected in high-energy γ rays (squares). Using their sample, they plotted γ-ray flux efficiency (y axis) against characteristic age (x axis). Apart from the 6 detected pulsars, only upper limits exist for the rest (triangles).

which leads to the absence of synchrotron radiation. However, because the trajectories of the particles at these heights are curved, the production of curvature radiation in the direction of motion becomes substantial. Sturrock stipulated that the high-energy γ-ray emission emanates from particle acceleration up to $h \approx R_{\rm pc}$ and is responsible for the observed GeV spectrum. However, such high-energy photons in a strong magnetic field of 10^{12} G are known to pair-produce according to $\gamma + B \longrightarrow e^- + e^+$; but not all do. It was assumed that half the γ rays produced this way escape. The rest — which pair-produce — generate particle cascades that give rise to the luminous, coherent radio emission. This emission is the curvature radiation from secondary $e^- - e^+$ pairs and can be explained by means of charged particle sheets: these are described as highly bunched particle-flow patterns, which leave the polar cap regions and transport charge into the outer magnetosphere. Provided these sheets overlap with each other as seen by an observer along the line-of-sight, there will be coherent radio emission.

This very early effort to explain the pulsed emission in a wide range of frequencies resulted in partial success. The fits to the first discovered pulsar, CP 1919, looked promising, and the predictions for the particle production based on this model favoured certain observations of the Crab nebula [233].

After the first steps towards a satisfactory description of pulsar emission had been made, the interest in further developing these ideas was high. So, in 1975, Ruderman and Sutherland tried to include in Sturrock's model the fact that the whole magnetosphere, including the open magnetic field lines, co-rotates with the star. While it is possible in Sturrock's model to explain accelerating potentials up to 10^{12} V, the addition of a co-rotating magnetosphere everywhere above the polar caps leads to accelerating potentials that are 100 times weaker, and so, as Ruderman and Sutherland argued, there cannot be enough acceleration to subsequently cause curvature radiation along the magnetic field lines. Instead, they envisaged a situation where the outflow of charge through the light cylinder is *not* compensated by the input of ions from the stellar surface. That led to a small charge gap above the polar caps, where everywhere inside $\boldsymbol{E} \cdot \boldsymbol{B} \neq 0$. Their reasoning was based on the neutron stars' surface temperatures, which are insufficiently high to allow the extraction of positive ions from the surface by means of an electrostatic pull against the ion binding energy. The estimated height of the gap at its full extent was $h_{\rm gap} \sim 10^4$ cm. However, the growth of such gaps stops when they reach this size due to $e^- - e^+$ cascades that are the result of the local-photon-field interaction with the huge magnetic field inside the gap. More specifically, the created pairs are accelerated in different directions, which causes the e^+ to exit the gap, whereas the e^- head towards the surface. However, because the gap formation allows for a strong potential difference to be built across it ($\sim 10^{12}$ V), the accelerated electrons produce curvature radiation that subsequently interacts with the magnetic field and, hence, produces even more particles. This whole process creates a spark along the gap, which equalises the potential between the edges and, consequently, restarts gap formation. The characteristic time between sparks was estimated to be $t_{\rm spark} \propto h_{\rm gap}/c \sim 10^{-6}$ s.

In Ruderman and Sutherland's model, the pulsar emission is mainly due to coherent curvature radiation from the e^+ that escape the gap. This explained the low-frequency emission from pulsars, i.e. radio and microwave, but it was unable to give reasons for the high-energy, X-ray and γ-ray emission. The model also tried to account for the observed double-peaked pulses (these had already been observed at that time from pulsars like the Crab and Vela) by describing the beamed geometry of the coherent radiation, as coming from nested cones that intersect the observer's line-of-sight. Interestingly, almost 10 years later, the foundations of this model were questioned in a paper by P. B. Jones, who argued that the cohesive energies of the lattice ions in the polar caps are too small to sustain a finite electric field boundary on the pulsar surface [234].

Sturrock's model, on the other hand, assumed that even hadrons can be accelerated in the polar cap zones and gave reasons for the observed high-energy spectrum. However, both models as well as variations of them have the location of the energy source in common: i.e. the polar cap region. This leads to the conclusion that both low-frequency and γ-ray emission should be observable for every pulsar, provided of course that the detection thresholds of the telescopes are low enough. The *CGRO* mission has already detected 6 high-energy pulsars that have previously been confirmed as radio pulsars (see Table 1.2). The particular case of Geminga, however, which appears very weak in radio frequencies, but which is amongst the brightest γ-ray sources, poses a challenge for the Polar Cap models (see section 3.2.2). Is Geminga unique, or is there differentiation between the emission mechanisms for different pulsars?

Polar Cap Simulations

Early computer simulations by A. K. Harding produced the formula

$$\dot{N}_\gamma(> 100\ \mathrm{MeV}) = 1.2 \times 10^{35} B_{12}^{0.95} P^{-1.7}\ \ \mathrm{ph\ s}^{-1} \tag{3.21}$$

which gives the γ-ray luminosity for Polar Cap–modelled pulsars [235]. It is worth noting that this spectrum was normalised to give the correct luminosity for the Crab pulsar. Harding's simulations did not deal with the particularities of particle acceleration in the polar cap zones, but instead they assumed a monoenergetic stream of particles being injected into the pulsar magnetosphere; various values of the magnetic field, $B_{12} = B/(10^{12}\ \mathrm{G})$, and period, P, were used. The models were in agreement with early observations of Vela pulsar and were well-hidden below the upper limits for other pulsars observed.

After the first simulations, there was need for expanding the analysis to further detail and including subsidiary contributions like the synchrotron emission from secondary $e^- - e^+$ pairs, as well as the acceleration of primary particles throughout the polar caps. From these simulations, Daugherty and Harding concluded that curvature γ rays at higher altitudes above the acceleration zones make substantial contributions to the final γ-ray spectra and allow the escaping photons to reach energies of 5 GeV [232]. In general, the deduced spectra have sharp cut-offs, a fact which can be assigned to the more efficient conversion of high-energy photons to e^-–e^+ pairs, in TeraGauss magnetic fields. Also, it was found that short-period pulsars produce softer spectra — i.e. with larger spectral indices ν, where $dN/dE \propto E^{-\nu}$ — than pulsars with longer periods. More recently, they refined their simulation code to allow particle acceleration up to several radii above the polar caps. Their model application to the Vela pulsar was able to reproduce the observed spectra and light curves [236].

In 1995, based on Harding and Daugherty's Polar Cap model, Nel and de Jager applied the following model fit to the EGRET data [237]:

$$\frac{dN_\gamma}{dE} = K \cdot \left(\frac{E}{1\ \mathrm{GeV}}\right)^{-\nu} \exp\left[-\left(\frac{E}{E_\mathrm{c}}\right)^{b}\right] \quad \mathrm{ph\ cm^{-2}\ s^{-1}\ GeV^{-1}} \tag{3.22}$$

which represents a single power law with a super-exponential cut-off. K is the normalisation flux at 1 GeV, and ν is the power-law index. These values depend on the observed spectra as measured with EGRET. On the other hand, the cut-off energy, E_c, and the super-exponential index, b, correspond to the position and

Pulsar	$K(\times 10^{-8})$ (ph cm^{-2} s^{-1} GeV^{-1})	ν	E_c (GeV)	b
Crab	24	2.08	30	2
Vela	138	1.62	8	1.7
B1706−44	20.5	2.1	40	2
B1951+32	3.8	1.74	40	2
Geminga	73	1.42	5	2.2
B1055−52	4	1.8	40	2

Table 3.2: (from [237] and [238]) The parameters of a power-law fit to the EGRET data, with a super-exponential cut-off, for the 6 EGRET pulsars.

steepness of the spectral cut-off, which is clearly evident only for the Vela pulsar and Geminga. For the rest of the EGRET pulsars, Nel and de Jager assumed $b = 2$, which is characteristic of spectra attenuated by the process of magnetic pair production. The derived values for the 6 EGRET pulsars are shown in Table 3.2.

In conclusion, it is worth saying that the Polar Cap models predict high-energy γ rays from pulsars, up to a few GeV. The deduced spectra are usually soft, with steep, super-exponential cut-offs (see for example [27]). They can account for the complexity of the observed pulses in terms of enhanced emission towards certain directions inside the conical beam. Various geometrical models have been devised to explain the observed pulses, which consider multiple components (i.e. bridge emission, interpulses, etc.). One of the first ones was the hollow-cone model by V. Radhakrishnan and D. J. Cooke [239], who considered that the emission lies inside a conical shell with finite width. Probably, such empirical models inspired theorists like Sturrock and others, who constructed detailed mathematical models in order to explain the observed pulsar characteristics. Similar empirical models, like the one by Rankin, for example, consider emission not only from a hollow cone but also from the central regions. Furthermore, the patchy-beam model by Lyne and Manchester considers a series of randomly distributed emission regions inside the conal beam. These models have become very popular, especially amongst radio astronomers [240],[241]. So far, the application of such models to γ-ray astronomy has been limited due to the relatively poor statistics compared to radio and optical observations. Nevertheless, as more sensitive experiments are being employed, the possibility of relying on these same models for an explanation of the observed γ-ray

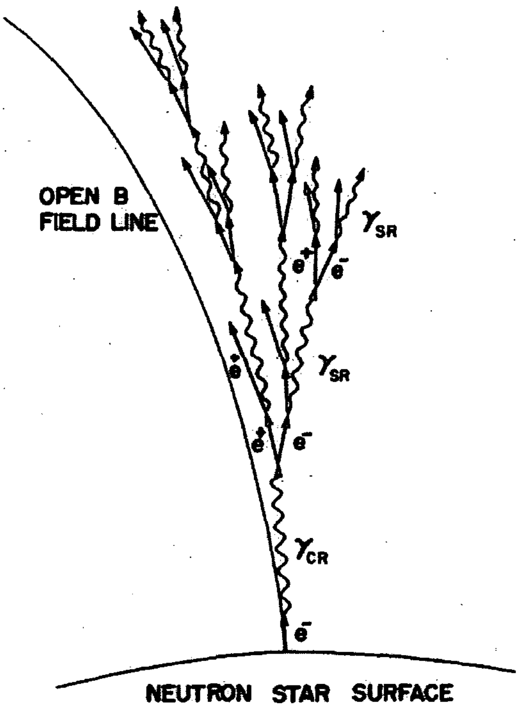


Figure 3.9: (from [232]) Schematic of a Polar Cap cascade. The primary particles are being accelerated above the polar caps and produce curvature γ-ray photons ($\gamma_{\rm CR}$). Further up in the magnetosphere, the same γ-ray photons meet the magnetic field lines at an angle and pair-produce. Secondary positrons are forced to spiral along the open magnetic field lines on their journey towards the light cylinder, while giving off synchrotron photons ($\gamma_{\rm SR}$).

pulses remains open.

The Outer Gap Model

In 1986, Cheng, Ho and Ruderman (hereafter CHR) proposed an alternative to the Polar Cap model, in order to explain the observed optical–γ-ray emission from young pulsars with short periods [243]. It was argued that under certain conditions the existence of charge-depleted areas within the light cylinder is possible. The areas that serve as electrostatic accelerating regions *are expected* to exist, according to the model, in the outer magnetosphere and are, therefore, called *outer magnetospheric gaps.* With simple arguments, the authors proved that these gaps cannot terminate within the light cylinder, but should have part of their boundaries outside it (see Fig. 3.5). Similarly to the Polar Cap model, these charge-free gaps serve as $e^- - e^+$

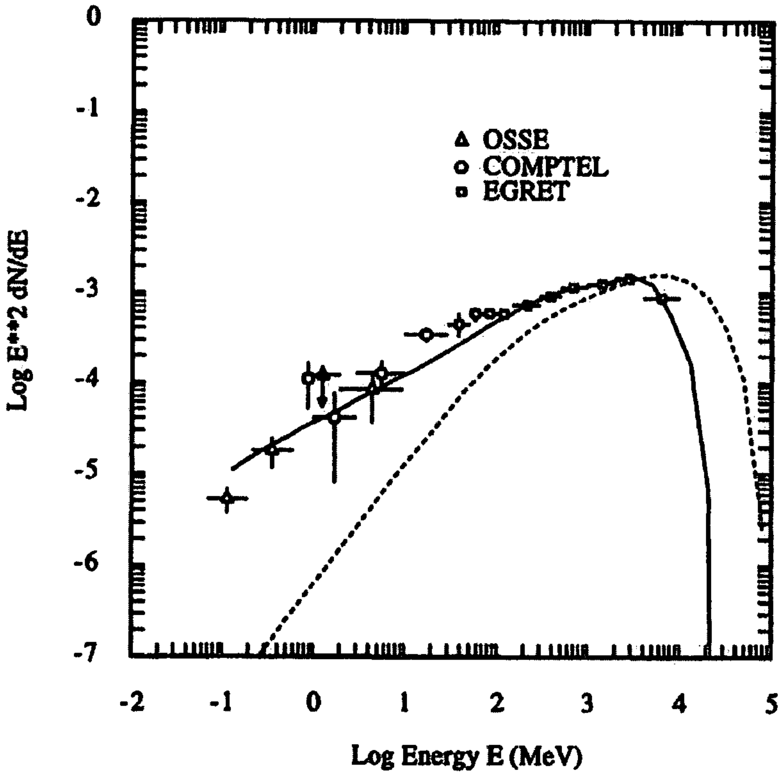


Figure 3.10: (from [242]) Typical Polar Cap spectral cut-offs at a few GeV energies. The dashed line represents the spectrum of the primary curvature radiation, and the continuous line is the full-cascade emission. Overlaid on the modelled spectra are data from three γ-ray experiments: EGRET, COMPTEL and OSSE.

pair acceleration regions, since the model stipulates that $\boldsymbol{E}\cdot\boldsymbol{B}$ does not vanish inside the gaps. However, the emission mechanisms in the Outer Gap model depend largely on the neutron star's properties and, as CHR described, differ between pulsars. The mechanisms involved are inverse Compton, γ–γ pair production, and curvature and synchrotron radiation: in other words, quite a rich diversity of processes. In the following, the principles of the Outer Gap model are described in more detail.

The existence of charge-depletion regions inside a pulsar's magnetosphere had been investigated before the formulation of the Outer Gap model [244]. All those models assumed aligned rotators with gaps forming well-within the light cylinder. CHR admitted that the existence of such gaps could come as a consequence of those previous models if non-aligned rotators are considered, or otherwise as the result of magnetospheric current flows between the star and the light cylinder. In their model, they adopted the latter view of the magnetospheric current flows, which they set forth *a priori*. The two postulates of the Outer Gap model are

- The huge electrostatic potential drop across the open magnetic field lines generates current flows of mainly non-charge-separated plasma, which runs through the charge-separated magnetosphere and drives particles away from the pulsar. This potential drop is proportional to $\Omega^2 B_{\rm s}$, and for pulsars which the Outer Gap model is interested in (i.e. young and fast-spinning, like the Crab), it can be as high as 10^{16} V.

- The return current (the flow of particles) that brings particles from outside the light cylinder, helping that way to close the circuit, runs through the null charge surfaces.

In their paper, CHR present a specific pattern for this current flow, which is shown in Fig. 3.11 for the case of non-aligned rotators. It is clear from this figure that the circuit closes outside the magnetosphere, but the answer to how this is done was left open.

We will see, now, how the two postulates of the Outer Gap model can lead to the creation of outer magnetospheric gaps. As a starting point, we assume that above the polar caps the charge-separated magnetosphere is positive: i.e. $\boldsymbol{\Omega}\cdot\boldsymbol{B} < 0$. In that case, the currents above the two polar caps are driving positive charge (non-charge-separated) towards the outer magnetosphere. This *outflow* of positive charge is balanced by the return current, which according to CHR has two components: first, a current of negative particles outflowing from the negatively charged inner magnetosphere and moving towards the light cylinder; and, second, an opposite *inflow* of positive particles coming from outside the light cylinder and moving towards the pulsar. The model makes the assumption that only the former has an important contribution.

If we imagine the dynamic magnetosphere of a non-aligned pulsar that tries to distribute the magnetospheric charge according to Fig. 3.11, then the return current which is responsible for the outflow of negative charge away from the null charge surface will leave a charge-depleted region (i.e. a gap) on one side of that surface (see Fig. 3.12). This gap is now electrostatically positive and repels the positive charges on the other side of the null charge surface (i.e. the starward side). This results in the expansion of the initial gap. As negative charge moves away from the null charge surface, the gap grows bigger and sends negative and positive charges to opposite directions. Naturally, the gap's growth will cease when the current flow is disrupted. This occurs when the potential difference across the gap balances the one that creates the current flow in the first place, which is the one that develops across the open magnetic field lines of the polar cap.

The description of such 'dead' pulsar models had already been given before the Outer Gap model, by Michel, Krause-Polstroff, Rylov, etc. [245],[244],[246]. However, the goal of the Outer Gap model was to provide a nursery for particle acceleration and high-energy emission in a *dynamic* magnetosphere: one that would sustain a continuous flow of charged particles. Hence, CHR argued that in the Outer Gap model the unrestrained growth of these gaps is held back by plasma creation, via pair production inside the gaps. Neutron stars with large $\Omega^2 B_s$ values should be capable of efficient acceleration of $e^- - e^+$ pairs to ultra-relativistic energies and subsequent production of high-energy photons via the various particle interactions. Such a scenario imposes serious constraints on the shapes and locations of the outer gaps:

> If one assumes that the charge-depleted regions are well within the light cylinder, then the large values of $\boldsymbol{E} \cdot \boldsymbol{B}$ inside the gaps will trigger such a high rate of pair production, which would soon fill the empty space with plasma. Clearly, such gaps cannot be long-lived.
>
> On the other hand, gaps developing in the outer magnetosphere *are* sustainable, but not anywhere. The copious amounts of accelerated charges along

> the open field lines will generally produce γ rays mainly via three production mechanisms: synchrotron, inverse Compton and curvature radiation. These γ rays are typically beamed along the path of the particles, and hence they propagate tangential to the magnetic field lines, which generates the so-called double fan-beam emission. Then, any gap formed on the convex side of an open magnetic field line would be a target for such γ rays and would unavoidably be quenched by pair production via γ–γ or γ–$\boldsymbol{B}$ interactions (see Fig. 3.13).
>
> Therefore, the only option for gap formation are places in the outer magnetosphere that are shielded by the γ-ray photon field. CHR defined those regions as long "slab gaps", with their concave border on the last closed field line and their convex border parallel to the open field lines, not too far above. This last border is shaped by the same γ rays which quench potential gaps higher in the magnetosphere.

Prior to γ-ray observations with instruments like EGRET, the predicted Outer Gap pulse shapes and spectra had to be calculated in order for the model to become competitive against Polar Cap models. CHR, in a follow-up publication after the original model, predicted Outer Gap spectra from optical to γ rays, for the Crab and Vela pulsars [247]. Unlike the Polar Cap models, which terminate γ-ray emission soon after entering the GeV regime, the Outer Gap model introduces an additional VHE component at TeV energies. As described in that paper, the high-energy spectrum is mainly due to secondary e^-–e^+ pairs that give off synchrotron photons, but also — for pulsars like Vela — inverse Compton γ rays from interaction with soft, IR, optical photons. The former emission is energetic enough to explain spectra up to a few GeV, but it is the inverse Compton photons that make up the TeV energies. However, according to CHR, not all pulsars can sustain a detectable TeV emission via the inverse Compton mechanism. For Crab-like pulsars, which have magnetic fields with relatively small radii of curvature, the primary e^-–e^+ pairs produce GeV curvature radiation which competes with inverse Compton for the dominant process. Therefore, TeV emission is expected to be low. However, they considered the possibility of fluctuations in the soft-photon production, which would allow a certain portion of the TeV emission to escape pair production and therefore boost the TeV signal. To support their argument CHR referred to early observations of the Crab pulsar from Dugway, which resulted in the detection of a weak, persistent

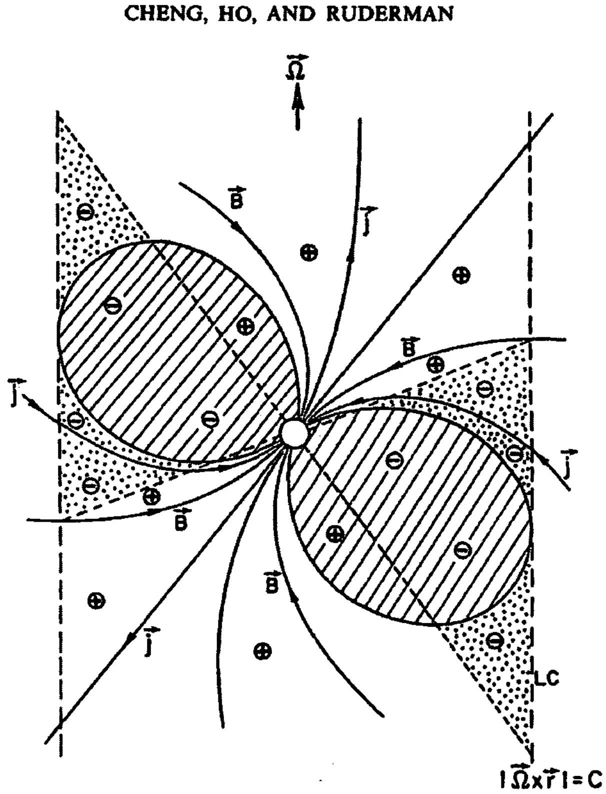


Figure 3.11: (from [243]) The configuration of a pulsar's magnetosphere according to CHR's Outer Gap model. Non-charge-separated plasma is accelerated along the open magnetic field lines from the polar caps to the light cylinder (positive charge in this case), and an equivalent return current from beyond the light cylinder flows towards the pulsar. The current direction in this figure corresponds to flow of positive charge; negative charge is outflowing in the dotted regions.

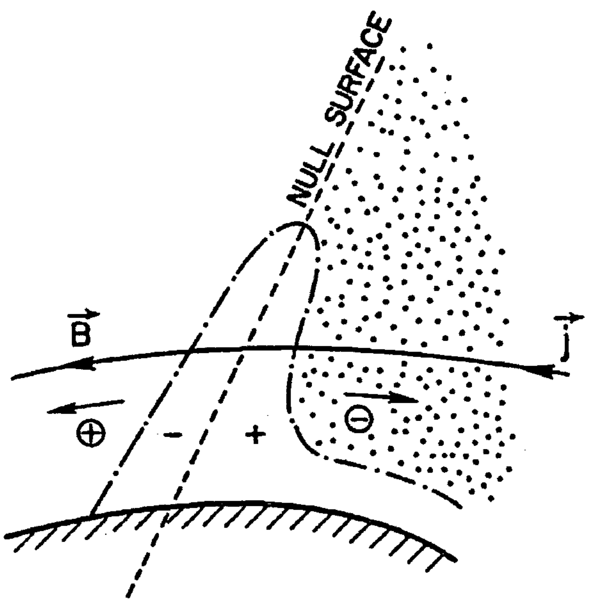


Figure 3.12: (from [243]) The formation and growth of charge-depleted regions (gaps) near the null charge surface. Assuming an outflow of negative charge transporting plasma away from the null charge surface — opposite to the direction of the magnetospheric current j — a charge-depleted region may be formed, whose electrostatic equivalent is a positively charged region — given the absence of negative charge. In this figure, the polarity of the charged plasma is indicated with the circled signs, "⊖" and "⊕", whereas the induced polarity of the charge-depleted regions — surrounded by a dashed–dotted line — with plain signs ("−" and "+"). The induced positive charge can repel negative plasma on the other side of the null charge surface and form a similar gap but with an opposite induced charge (negative). As more plasma leaves the region near the gap (i.e. the dot-filled region), both negative and positive charge-depleted regions grow in size. This process is naturally halted when the potential difference across the gap is high enough to balance that which drives the non-charge-separated current of positive and negative plasma.

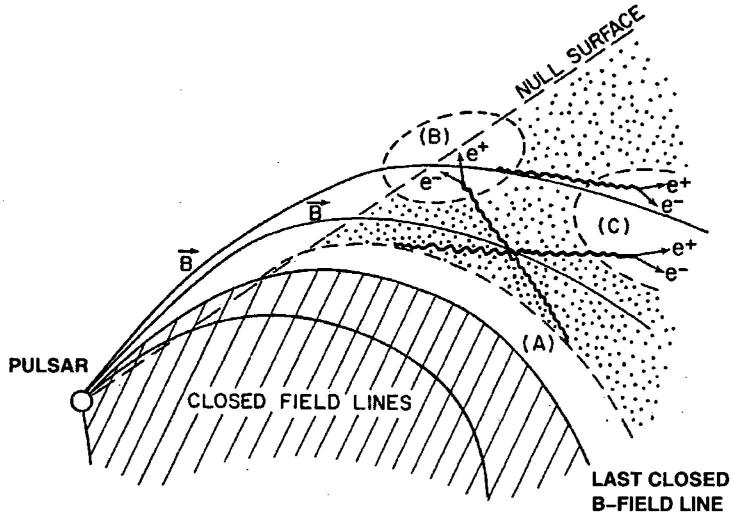


Figure 3.13: (from [243]) Different scenarios of gap formation in the outer magnetosphere. Gaps (A) and (C) do not terminate within the light cylinder. Gaps (B) and (C) are on the convex side of open magnetic field lines, and hence they are targets for γ rays produced tangential to those lines. Consequently, these gaps are quenched by pair production within their volume. However, long "slab gaps" like (A) are shielded by such γ rays because they border with the last closed field line; gaps of type (A) are the basis of the Outer Gap model.

periodic signal in the TeV range [93].

On the other hand, slower rotators like the Vela pulsar, with substantially larger radii of curvature, are expected to have bright inverse Compton components. Moreover, the weak outer-magnetospheric magnetic field of the Vela pulsar ($\sim 10^4$ G) — as opposed to the Crab pulsar's ($\sim 10^5$–10^6 G) — makes it possible for most of the TeV γ rays to escape without pair-producing on the field lines. Such pulsed TeV emission was observed from the Vela pulsar by Bhat *et al.*, in 1979 [248].

Following CHR's work, an important effort to model the observed complexity of the γ-ray pulses was made by Chiang and Romani [249]. They investigated the γ-ray emission from a 3-dimensional magnetosphere of an Outer Gap accelerator, for different viewing angles. Initially, they assumed uniform emission from the outer gaps along the last closed field lines and did not account for the spatial dependence of emissivity across the magnetosphere.

Further work based on the Outer Gap scenario was carried out by Romani [250], who provided the formulation for the γ-ray efficiencies from Outer Gap accelerators and showed how they evolve with time. In addition, Romani dealt with the phase-resolved spectra for Vela-like pulsars and succeeded in fitting Outer Gap–model functions to the observed fluxes, for the individual pulse components. However, at the highest available energies, the observed data points have large errors, which renders the validity of the models uncertain.

In 1997, Zhang and Cheng (hereafter ZC) expanded the classical view of the Outer Gap model and went on to divide the different pulsars into three classes, according to the size of their outer gaps [251]. For pulsars like Geminga and PSR B1055$-$52, the outer gaps occupy a large fraction of the outer magnetosphere ($\approx$ 70%), which ZC characterised as thick gaps. Crab-like pulsars, on the other hand, retained the classical view of pulsars with thin slab gaps. The intermediate class, with medium-sized outer gaps, includes pulsars like PSR B1706$-$44. ZC stipulated that the high-energy emission for thick-gap accelerators comes from primary $e^- - e^+$ pairs rather than secondary, which is also the case for thin outer gaps. Based on that, they derived the X-ray and γ-ray spectra for the above pulsars, which matched well the available OSSE, COMPTEL and EGRET observations in the 100 MeV–10 GeV energy range and was consistent with the corresponding upper limits at lower energies. Unfortunately, the difference between the thick- and thin-gap scenarios for the Crab pulsar and PSR B1055$-$52 is comparable to the EGRET error bars and is well below the COMPTEL upper limits. Hence, contemporary observations are unable to discriminate between the alternative Outer Gap scenarios (see Fig. 3.14).

Outer Gap Simulations

More recently, a very important contribution to the Outer Gap model prediction was made by Hirotani and Shibata (hereafter HS). Their work demonstrated the existence of an Outer Gap accelerator as a solution of the equations that govern plasma electrodynamics, and the γ-ray spectra for the Crab, Vela, PSR B1951+32 and other pulsars resulted consequently. The set of equations that describe the one-dimensional structure of an electrodynamic accelerator along the field lines is called the Vlasov equations, and it is a combination of the one-dimensional approximation of the Poisson equation for the charge distribution, the Boltzmann equations for the γ-ray distribution, and the particle continuity equations for positrons and electrons. The Poisson equation defines the accelerating electric field parallel to the magnetic field lines; the particle continuity equations describe the production of $e^{\pm}$ inside the gaps; and finally, the Boltzmann equations describe the γ-ray production from accelerated $e^{\pm}$. This set is provided here for reference, but a detailed explanation can be found in [252],[253],[254]:

- Poisson equation

$$\frac{dE_{\|}}{ds} = 4\pi e \left(N_{+} - N_{-} + \frac{\Omega B_z}{2\pi c e} \right) \tag{3.23}$$

- Continuity equations

$$\pm cB \frac{d}{ds} \left(\frac{N_{\pm}}{B(s)} \right) = \int_0^{\infty} d\epsilon_{\gamma} \left(\eta_{p+} G_{+} + \eta_{p-} G_{-} \right) \tag{3.24}$$

- Boltzmann equations

$$\pm cB \frac{\partial}{\partial s} \left[\frac{G_{\pm}(s, \epsilon_{\gamma})}{B(s)} \right] = -\eta_{p\pm} G_{\pm} + \eta_{c} N_{\pm} \tag{3.25}$$

where $E_{\|}$ is the accelerating electric field, parallel to the local B-field; s is the distance along the field lines; $N_{\pm}$ is the number density of positrons and electrons, respectively; $\mathbf{\Omega}$ is the angular velocity of the neutron star; B_z is the magnitude of the magnetic field component parallel to $\mathbf{\Omega}$; ϵ_{γ} is the γ-ray energy normalised to the electron rest energy; $\eta_{p\pm}$ is the pair production rate per γ ray; $G_{\pm}$ is the distribution

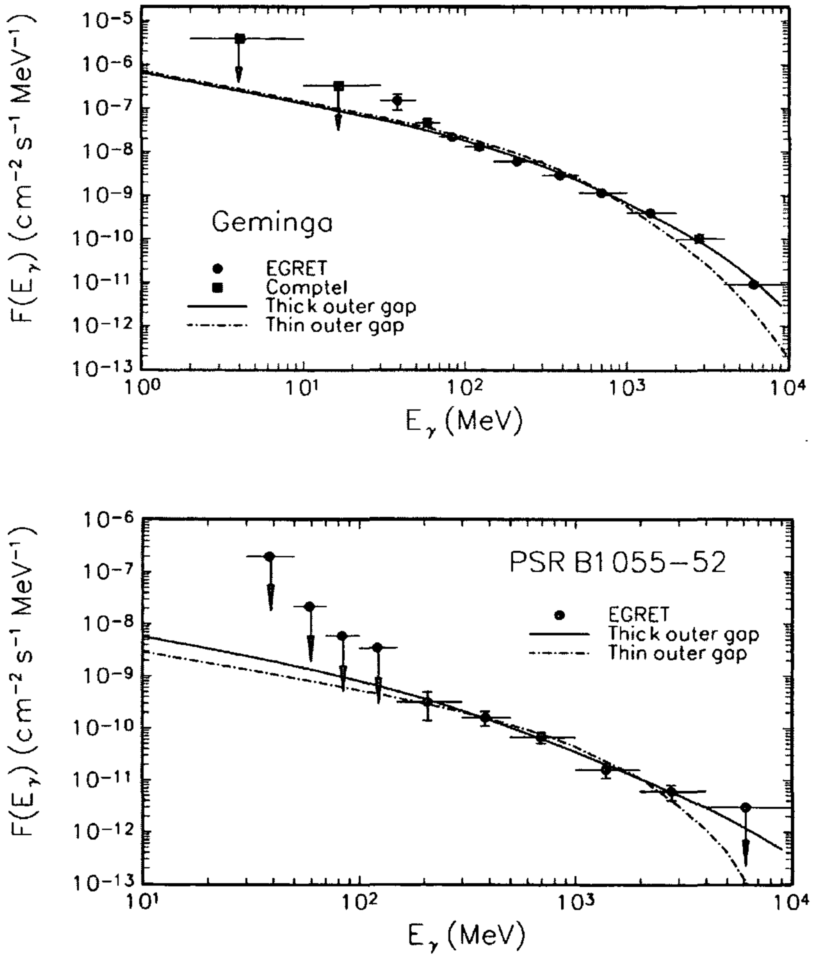


Figure 3.14: (from [251]) The spectral fit of the EGRET and COMPTEL data for Geminga and PSR B1055−52, with the two possible Outer Gap scenarios: the thick (solid lines) and thin (dashed–dotted lines) gap. According to ZC, both the Geminga and PSR B1055−52 are expected to have thick gaps in their outer magnetospheres.

function for inwards (−) and outwards (+) propagating γ rays; and, finally, η_c is the curvature radiation rate per particle.

HS solved Eq. 3.23–3.25 numerically by defining the appropriate boundary conditions. This way, they provided a self-consistent, quantitative solution for the Outer Gap model, calculated the critical quantities $E_{\|}$ and ΔV_{gap} and derived γ-ray spectra from first principles. Their results show that the accelerating fields across the outer gaps can reach values ($\sim 10^8$ V m^{-1}) that are higher than those previously assumed. However, this updated value of the electric field affects mainly the γ-ray luminosity due to curvature radiation, which is sensitive to $E_{\|}$, and not the luminosity from inverse Compton, which is responsible for the TeV component of the spectrum. The latter is sensitive to the size of the gaps and the low-energy photon flux. Hence, for narrow gaps, it was calculated that the GeV luminosity is 2–3 orders of magnitude higher than its TeV counterpart. The Crab pulsar's spectrum which was derived from HS's calculations is shown in Fig. 3.15, where one can see that, under certain boundary conditions for the electric field, the synchrotron component can be extended up to tens of GeV, and that there is also a low luminosity TeV component; for more details see [252]. The TeV spectrum in this plot was derived from an IR component that came from a single power-law interpolation between the radio and optical spectra.

The Polar Cap and Outer Gap models are strong candidates for a detailed description of high-energy emission from pulsars. The key criterion in the determination of the correctness of either model is the reproduction of the observed pulse shapes and spectra. Nevertheless, the scientific community is open-minded, and alternative approaches might be proven more successful when the present or future experiments collect the necessary data to constrain theoretical predictions. There have already been more recent attempts to explain the observations, which look promising. Models like the *two-pole caustic* model [258], which assumes a magnetospheric gap extending from the polar cap regions to the light cylinder, and the *Deutsch field* γ-ray pulsar model [259], which assumes a magnetosphere with low charge density and treats the pulsar as a highly conductive spinning sphere, are just two of the latest developments in theoretical pulsar astrophysics. There is no doubt that there will be even more in the future. The real question is, however, *which* one will be established and, furthermore, if there is indeed a single model that describes pulsars, or if, instead, they are objects with a diversity of underlying emission mechanisms.

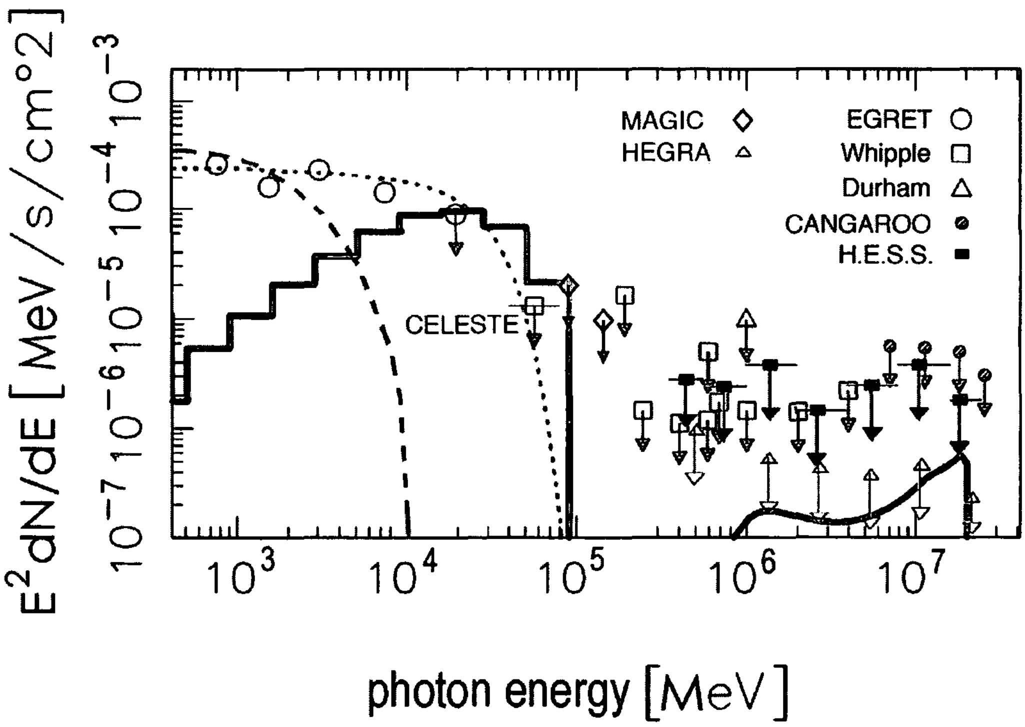


Figure 3.15: Crab's pulsed γ-ray spectrum (solid lines), as was calculated by HS, for a magnetosphere with magnetic inclination $\alpha_\text{i} = 45°$ and an IR component which comes from the single-power-law interpolation between the radio and optical bands. For comparison the plot includes the Polar Cap prediction for the Crab pulsar (dashed line) and a power law with super-exponential cut-off fit (grey, dotted line) [255],[238]. Also shown in this plot are the upper limits on the pulsed flux from different experiments [252],[256],[257].

3.2 Pulsars in their Natural Habitats

3.2.1 Introduction

The image of an isolated neutron star that spins together with its magnetosphere in vacuum might be intriguing for a theorist. But observations show that most pulsars are part of either plerion or binary systems. Naturally, all pulsars are born in supernova explosions. Surprisingly, there are only a few clear associations with plerions apart from the Crab and Vela pulsars (see e.g. [260]). Binary and millisecond pulsars are also found in the Galaxy, and observations show that the vast majority of them are accompanied by main-sequence or giant star companions. Their non-degenerate companions act as beacons, which makes these systems more easily observable. Double-pulsar systems are also thought to be relatively abundant. Examples of such objects are the binary system PSR B1259−63/SS 2883, a 50-ms pulsar orbiting a Be star, and the second fastest-rotating pulsar known, PSR B1937+21, with a remarkable 1.55-ms period [261]!

Despite the mere handful of known nebulae hosting pulsars, observers persistently search for evidence which will verify the models of stellar evolution. However, it is not always easy. The case of PSR B1706−44, an isolated pulsar that does not show evidence for a connection with a surrounding nebula, regardless of the copious observational efforts, is fairly typical. A favourite scenario, and quite a reasonable assumption, is that in most cases the ejected supernova material has faded to a degree which has rendered it undetectable. However, not all pulsars fall into this category. Young pulsars, like PSR B1706−44, must have undergone a different evolutionary scenario, or perhaps they were displaced by the supernova explosion [72].

Similarly, the observed population of binary pulsars falls short of expectations: the number of binary star systems in the Galaxy suggests that the corresponding binary pulsar population should be higher. Again, it is very likely that the catastrophic events at the end of the pulsar companions' lives could have broken up these systems. The discovery of isolated millisecond pulsars supports this scenario, but only on the basis of the current belief that these pulsars were spun up to millisecond periods by an accretion process from a companion star. A good source of information on the origin of millisecond pulsars is [262], where supportive arguments and open issues are addressed. However, such a scenario requires that the millisecond pulsar's magnetic field is low ($\sim 10^8$ G), otherwise the dipole radiation losses (see

Eq. 3.10) would slow them down to periods of ~ 1 s. An estimate of the characteristic time, after which that would occur, can be calculated for millisecond pulsars from Eq. 3.11:

> After rearranging the terms and invoking the characteristic age, τ, through the relation $P\dot{P} = P^2/(2\tau)$, we get
>
> $$\tau = \frac{3c^3 I}{16\pi^2 R_{\rm ns}^6} \cdot \left(\frac{P}{B_{\rm s}}\right)^2 \tag{3.26}$$
>
> For a millisecond pulsar with a TeraGauss magnetic field, this equation yields a mere age of ≈ 16 y, which is too short for any millisecond pulsar distribution to be observable. On the other hand, a magnetic field of $\sim 10^8$ G gives a characteristic age of 10^9 y, which would lead to a statistically significant sample of millisecond pulsars from contemporary observations.

As in the case of normal pulsars, almost all of the millisecond pulsars' rotational energy is carried away by the wind of accelerated particles: the pulsar wind. Nevertheless, a small fraction is lost via high-energy emission. Amongst the radiation mechanisms with which millisecond pulsars lose energy is that described by the model of Zhang and Cheng [263]: they assumed an Outer Gap–type magnetospheric configuration, from which X-rays arise (a) via synchrotron emission from the return current of particles; (b) via thermal emission from the polar caps due to inwards propagating curvature photons that are produced in the strong local magnetic field, close to the pulsar's surface (soft, thermal X-rays are produced this way); and, finally, (c) via thermal emission due to heating of the polar caps by the return current of particles (medium–hard, thermal X-rays). Also, in such scenario, γ rays are generated in the outer gaps via curvature radiation emitted by the return current of particles.

Alternatives to the conventional evolutionary track of millisecond pulsars have also been proposed: e.g. the possibility that close binary systems containing a white dwarf and a neutron star have evolved into a single millisecond pulsar via coalescence of the two objects; in such model, gravitational radiation losses force the two objects to merge into a single one [264].

In the following paragraphs, the macroscopic properties of pulsars as they appear from an observer's point of view are discussed: i.e. their population distribution; their companions; and their immediate surroundings beyond their magnetospheres, where the information we have on them is evidence from direct observations rather than theoretical modelling. We will focus on high-energy pulsars, which are the subject of this thesis. However, the statistics at those energies are limited, but also of the highest interest. On the contrary, the radio observations of pulsars have flourished in the past years, and extensive bibliography is available. For such information the reader is referred to [184],[265] and [266].

3.2.2 The Milky Way Family

So far, all the observations of pulsars have been confined within the Galaxy, including the two Magellanic Clouds, LMC and SMC. Extragalactic pulsars are hard to detect — even in radio — since their apparent luminosity at the Earth would be very faint. Hence, γ-ray detections from such distances seem almost impossible. But it is not only distance that constrains the high-energy observations of pulsars. The statistical sample is heavily reduced by the fact that, from a fixed position such as the Earth, one can hope to detect only the fraction of pulsars whose beams happen to cross the Earth's path. In addition, a further bias is introduced by pulsars with short periods and low luminosities, which are harder to detect than slow and bright ones. An elegant approach towards correcting for these biases was made by Phinney and Blandford [197]; see also Vivekanand and Narayan [267]. Using observations, they constructed the scaling factor

$$\xi(P,L) = \frac{\int\int_{\text{Galaxy}} \Sigma(R,z) R dR dz}{\int\int_{P,L} \Sigma(R,z) R dR dz} \tag{3.27}$$

which is the ratio of the volume (in cylindrical co-ordinates) that the whole Galactic distribution of pulsars occupies, based on an assumed pulsar density function $\Sigma(R,z)$, to the volume in which a specific pulsar with period P and luminosity L could be detected by the various surveys. Note that $\xi(P,L)$ is primarily a function of pulsar luminosity and period. In order to correct for the biases, pulsars with low luminosities and short periods are assigned small detectable volumes in the above calculation. In practice, the above scaling factor can be calculated for every detected pulsar, given an amount of Monte Carlo simulations which use the Galactic volume that has been explored by the major pulsar-detection surveys. Hence, if ξ_i is the value of this ratio for a pulsar whose luminosity exceeds the detection threshold, then the estimated number of active pulsars in the Galaxy is given by

$$\sum_{i=1}^{N_{\rm obs}} \frac{\xi_i}{f_i} \tag{3.28}$$

where $N_{\rm obs}$ is the detected number of pulsars from surveys, and $f_i = \Omega_i/(4\pi)$ are the beaming factors of the $N_{\rm obs}$ pulsars: i.e. the fraction of the sky that the pulsar beam subtends.

Unfortunately, as was first noted by Kalogera [268], the above scaling factors are representative of the entire pulsar population, *only* if the number of observed pulsars, $N_{\rm obs}$, exceeds 10; otherwise, a small-number bias is introduced. This means that if the 7 γ-ray pulsars detected so far could be regarded as a separate population from the low-energy radio pulsars, the above technique would have had difficulties in estimating the total number of high-energy pulsars in our Galaxy. Fortunately, the present consensus converges towards the expectation that rotation-powered radio pulsars with typical magnetic fields of $\sim 10^{12}$ G should have a γ-ray component, since both Outer Gap and Polar Cap models allow it. A second reason comes from pulsar energetics: the spin-down luminosity of the observed radio pulsars, in conjunction with their radio spectrum, suggests that only a small fraction of their energy output is in radio frequencies. The rest is distributed over the optical–γ-ray frequency range. Fig. 3.16 highlights the fact that the 7 high-energy pulsars detected so far have a power spectrum which peaks at γ rays. Moreover, they are all ranked at the top of the list in terms of observability, $\dot{E}_{\rm rot}/d^2$ (d is the distance to the pulsar), as Fig. 3.17 clearly shows.

Not all pulsars are rotation-powered, however. While normal and millisecond pulsars have magnetic fields $B_{\rm s} \lesssim 10^{14}$ G, with the latter type's fields being as weak as $B_{\rm s} \sim 10^8$ G, there is a small fraction of the discovered sample which has fields in excess of $\sim 10^{14}$ G. This sample is represented by the pulsars in the upper right corner of the P–$\dot{P}$ diagram (see Fig. 3.19); their ultra-strong magnetic fields has given them the name *magnetars*. The dominant process of emission from magnetars is not instigated by rotation but by magnetic-field decay. The strong magnetic fields of the magnetars make the process of energy loss through the pulsar wind a lot more efficient, which causes them to spin down much more rapidly than normal pulsars; also, kinetic energy is lost through the emission of magnetic waves. Hence, magnetars quickly become (after $\sim 10^4$–10^5 y) slow rotators, as their position in the P–$\dot{P}$ diagram reveals. Furthermore, the enormous magnetic forces exerted on the neutron star's interior and surface generate frictional heat by shifting stellar mate-

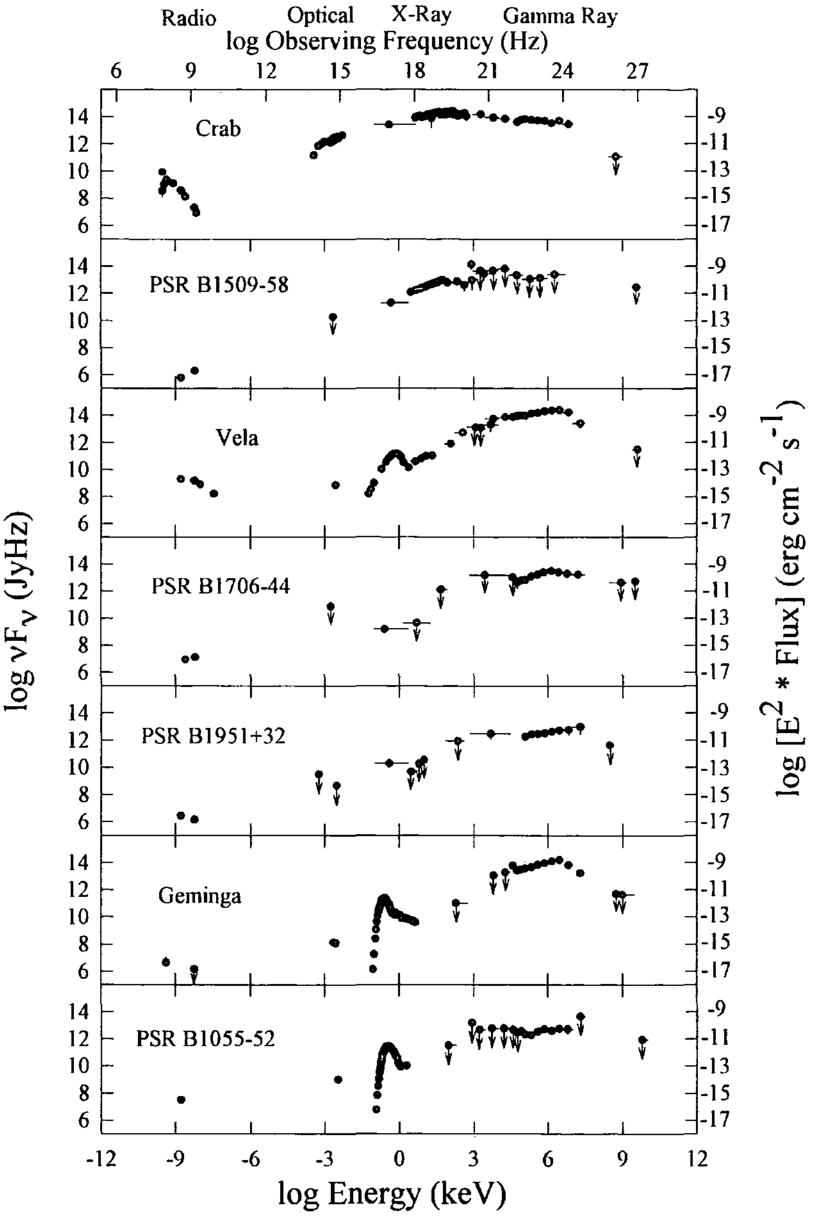


Figure 3.16: (from [5]) Multi-wavelength spectra for the 7 γ-ray pulsars. Amongst others, these plots show that the energy spectrum peaks at γ-ray energies, and that the corresponding radio spectrum is much lower in comparison. This hints that most of the rotation-powered pulsars should have a γ-ray component. However, this component seems to terminate in the VHE range, as only upper limits exist beyond 50 GeV. It should be recognised, nevertheless, that the reason for the poor sample of γ-ray pulsars may be the limitations of our detectors rather than the lack of sources in the Galaxy.

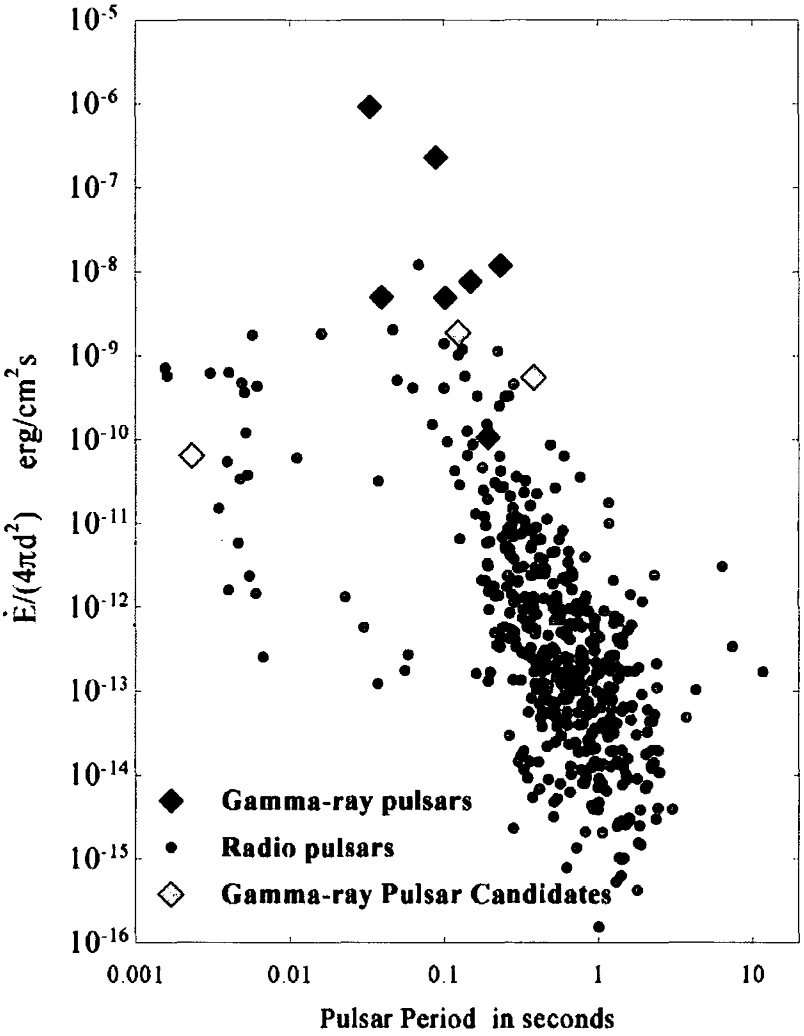


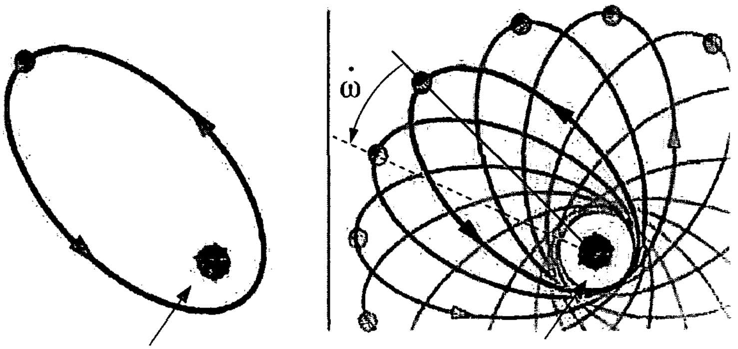
Figure 3.17: (from [269]) Observability plot for the radio (black dots) and γ-ray (diamonds) pulsars. When ranked according to spin-down flux, $\dot{E}_{\rm rot}/d^2$, 6 out of the 7 detected γ-ray pulsars (black diamonds) come at the top of the list amongst the pulsar population. This does not necessarily imply that all γ-ray pulsars should follow the same trend, since the available sample is very limited; it could also mean that our experiments have only been able to observe pulsars which are most efficient in converting rotational energy into γ rays (for a discussion, see [27]). Also in this plot there are 3 unidentified EGRET sources (light blue diamonds) that are believed to be high-energy pulsars.

rial; this is more true in the first $\sim 10^4$ y of a magnetar's evolution, during which period magnetars are observable as bright X-ray sources. Magnetars have been connected with the sudden bursts of hard X-ray emission that have been observed from Soft Gamma-ray Repeaters (SGRs) and with the abnormally high X-ray luminosity (10^{34}–10^{36} erg s^{-1}) of Anomalous X-ray Pulsars (AXPs): the bursting behaviour of magnetars is thought to arise from the occasional fractures that the neutron star's crust sustains under the huge magnetic forces [270],[271].

In the diagonally opposite to the magnetars corner of the P–$\dot{P}$ diagram are millisecond pulsars. Despite the expectation of finding these objects exclusively in binary systems, roughly 20% of the discovered sample is isolated. Some of the proposed evolutionary scenarios leading to isolated millisecond pulsars consider the disruption of the binary systems by means of (a) a supernova explosion of the pulsar's companion (this is most relevant in HMXRBs, where the post–main sequence star is massive enough to become a supernova); and (b) the vaporisation of the stellar companion through its interaction with the pulsar's wind (most relevant in LMXRBs). Examples of millisecond pulsars in the process of vaporisation are PSR B1957+20 and PSR B1744–24A [272],[273].

Occasionally, HMXRBs survive disruption and a system of two pulsars orbiting each other can be formed. An example of such an exotic binary is PSR B1913+16, discovered with Arecibo in 1974 [274]. According to general relativity, the strong gravitational field due to the presence of neutron stars in this system curves spacetime in their vicinity. As a result, the periastron position shifts monotonically with time (Fig. 3.18). This effect is also evident — on a smaller scale — in Mercury's orbit around the Sun. The phenomenon is called relativistic precession and should not be confused with the Earth's axial precession every 26,000 y: the latter is the result of the gravitational torques between the Earth and the Sun, which try to align the axes of the two objects. The amount of relativistic precession in PSR B1913+16 is 4° y^{-1}, which makes its companion almost certainly a neutron star.

At present, the most relativistic double-pulsar system known is that of PSR J0737–3039A/B, with $\dot{\omega} \approx 17^\circ$ y^{-1} [275]. The system exhibits a 'tight' orbital configuration (the tightest known) with $e = 0.088$, which results in a revolution period of only 2.4 h. Given their mutual proximity and the high rate of periastron advance, this system is expected to merge via gravitational emission into a single object, in a timescale of 85×10^6 y. More about this exotic configuration is mentioned in section 6.2.



(a) low–mass star

(b) high–mass/compact star

Figure 3.18: (from [276]) The phenomenon of relativistic precession. The gravitational field of a star creates a curved spacetime around it. A planet orbiting this star is bound to follow the orbital paths determined by the curved spacetime. The result is a monotonical shift, i.e. a precession, of the periastron position with time. (a) A planet's precession around a star is not apparent for low-mass stars, in which case the orbital orientation remains practically unaltered. (b) However, massive stars and compact objects, like white dwarfs and neutron stars, curve spacetime substantially, which leads to precession rates, $\dot{\omega}$, as high as several deg y^{-1}.

The Magnificent Seven

We shall now describe, in brief, the 7 high-confidence detections of γ-ray pulsars found so far in the Milky Way. The currently very limited sample of resolved γ-ray pulsars gives us this unique opportunity to review them individually. Nevertheless, future experiments, like *GLAST*, give an estimate of $\sim$ 500–1,000 potential detections by the end of their operation.

All the discovered γ-ray pulsars are young, with characteristic ages ranging from 10^3 to 10^6 y. According to their derived properties — which were based on the standard model — they exhibit high surface magnetic fields and high accelerating potentials across their open magnetic field lines (see Fig. 3.19) — a fact which makes them very energetic and potentially detectable from Earth. The EGRET instrument on-board *CGRO* was the first experiment to reveal high-energy ($>$ 100 MeV) pulsed emission from at least 4 previously undetected γ-ray pulsars: PSR B1706−44, PSR B1951+32, Geminga and PSR B1055−52; although many of the unidentified EGRET sources may well be proven to be HE pulsars, too (see section 1.2.3). As was mentioned in chapter 1, the Crab and Vela pulsars had already

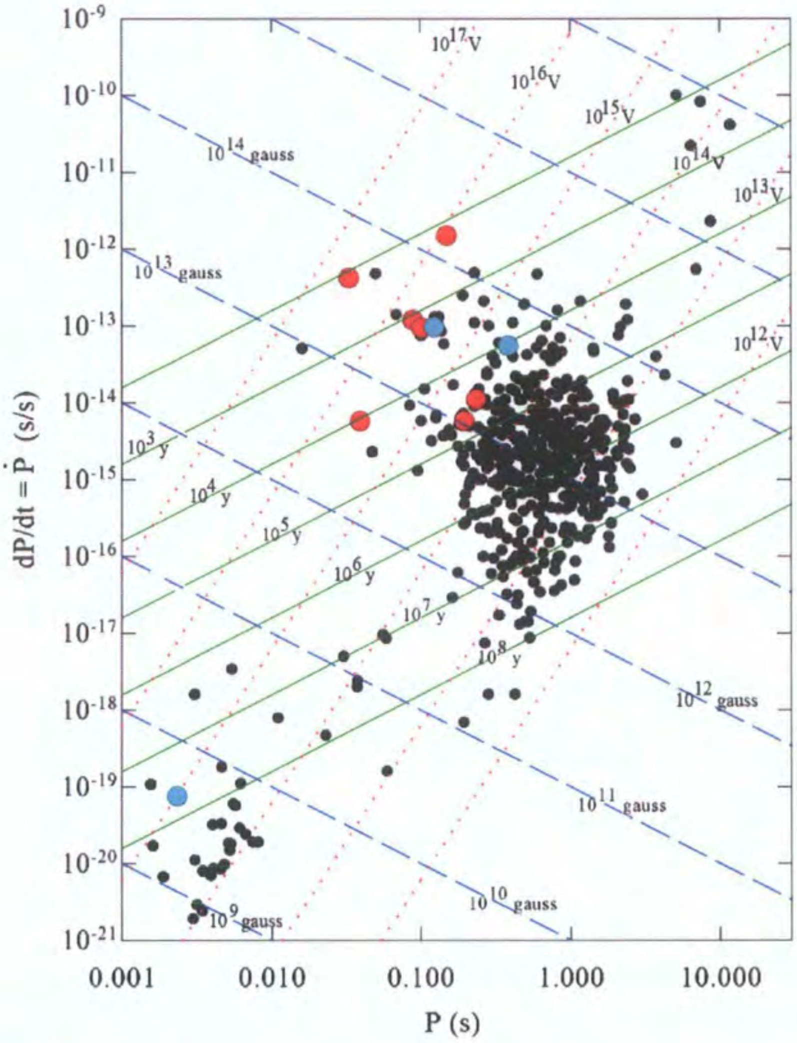


Figure 3.19: (from [269]) The population of radio and γ-ray pulsars on a P–$\dot{P}$ diagram. The black dots are pulsars from which no γ-ray emission has been detected. The 7 high-confidence γ-ray pulsar detections are represented by the red circles, and another 3 lower confidence γ-ray detections, by the light blue circles. The set of dotted, red lines represents the potential difference across the open magnetic field lines, for rotation-powered pulsars. The set of dashed, blue lines divides the pulsar population according to surface magnetic field, $B_{\rm s}$, and the set of solid, green lines, according to spin-down age. Along with any biases that γ-ray pulsar observations might introduce, it seems that γ-ray pulsars are characterised by large surface magnetic fields and high accelerating potentials.

been established as HE emitters with the observations of *COS-B* in the 70s [277]. On the other hand, PSR B1509−58 was only detected up to 30 MeV with COMPTEL. None of the above objects' pulsed signatures have been detected beyond EGRET's range ($\gtrsim$ 20 GeV) with VHE experiments, which implies that there is an energy cut-off for all these pulsars somewhere in the unexplored energy range (see Fig. 3.16). As mentioned in section 3.1, the competing theoretical models predict such a behaviour, but they also differ in the prediction of the cut-off energies. The answer to which model explains the observations better is in the hands of ground-based experiments, like H.E.S.S. and MAGIC, which probe these objects at high energies, or satellite experiments, like *GLAST*, which will try to extend EGRET's energy range and find the cut-off energies.

As one can expect, the locations where all the high-energy pulsars were discovered lie close to the Galactic plane. Previous surveys with Explorer XI and *COS-B* had revealed the prominence of the Galactic plane as the strongest γ-ray source in the Galaxy, something that was also verified by the EGRET all-sky survey. This survey brought to light numerous Galactic γ-ray sources. Unfortunately, most of them were unidentified due to the instrument's spatial resolution, which made it impossible to correlate them with known radio sources (see Fig. 3.20). However, if some of these sources are proven to be radio-quiet γ-ray pulsars, like Geminga, then the statistics alone will allow us to estimate their total number in the Galaxy. A discussion on this possibility can be found in [278], where the conclusion is largely based on the observed EGRET fluxes in relation to the assumed rate of supernova occurrence in the Galaxy. On average, these fluxes are higher than those observed from the high-confidence pulsars, which indicates either a population of very young and distant, energetic pulsars or low-luminosity, nearby ones — or, of course, objects of a different nature. If these objects are pulsars, then the numerous sources detected along the Galactic plane, where EGRET's sensitivity is lower due to the high background, must all be at least as young and energetic as the Crab, PSR B1706−44, etc. This is, however, something that is hardly supported by the observed supernova birth rate. On the other hand, the smaller sample of high-latitude, unidentified sources, which are at closer distances, is more likely to consist of pulsar candidates, with potentially low luminosities.

Fig. 3.21 shows the observed pulse shapes of the 7 high-energy pulsars at different wavelengths. It can be seen that in all cases, apart from the special case of the Crab, the transition from the radio band to the optical waves completely changes the

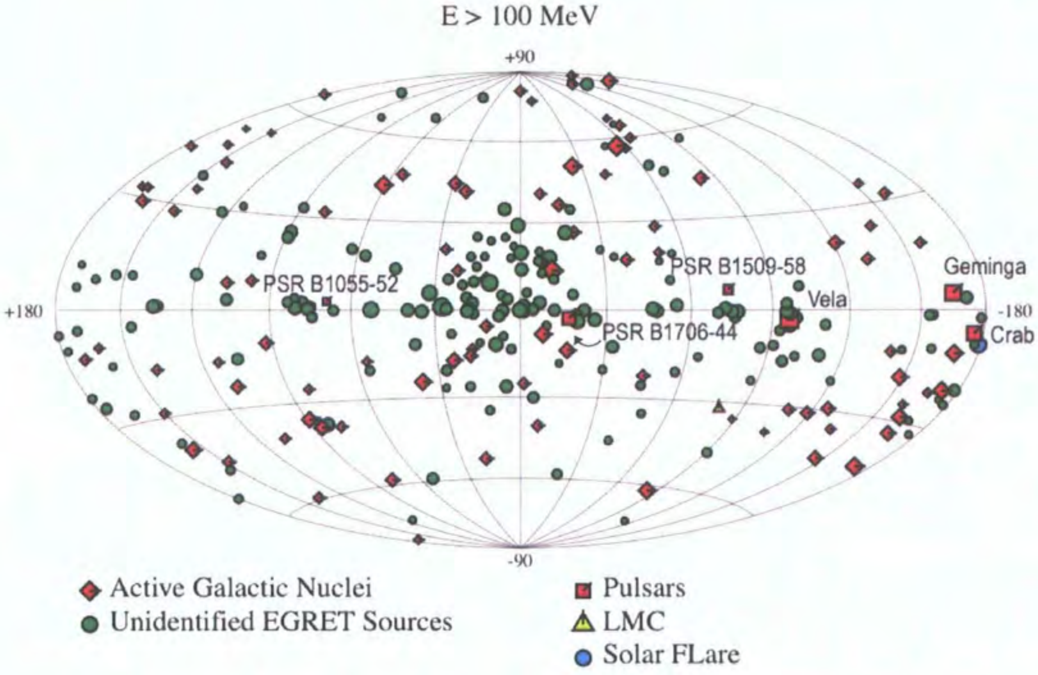


Figure 3.20: (from [14]) The Hammer–Aitoff map of the 3rd EGRET catalogue sources. Amongst others, this plot shows the positions of the 6 γ-ray pulsars detected at a high confidence close to the Galactic plane (squares). The size of the symbols in each set denotes the magnitude of the detection's highest intensity. It is also apparent that a large fraction of the detected sources could not be identified with any previously studied source (circles). Potentially, these sources might be proven to be pulsars, but the probability seems low, especially for the ones along the Galactic plane (see text).

phase location and the shape of the profile's components; and there is a significant change in pulse morphology, even between soft and hard X-ray observations. So far Geminga's case seems unique, as it has not been detected at a high confidence in radio waves. Nevertheless, what is of most interest for VHE γ-ray astronomy is the top end of the energy spectrum. At the bottom row of Fig. 3.21 there are six histograms of the most energetic pulses that EGRET has managed to detect (> 5 GeV). The most interesting feature is the suppression of the trailing component and the prominence of the second peak, which now dominates all profiles except PSR B1706−44's. However, this pulsar's profile can be considered broad and single-peaked, with duty cycle $\delta \approx 0.3$. The rest of the pulsars display the same behaviour at multi-GeV energies.

Geminga (PSR J0633+1746) is located in the constellation Gemini, and despite being the closest, it is perhaps the most enigmatic γ-ray pulsar yet. It rotates once every 237 ms and is a relatively young neutron star with $log(\tau/y) \approx 5.5$. Its detection in γ rays was first achieved by the *SAS 2* satellite and was followed by extensive observations with *COS-B* [281],[282]. Subsequent observations in the X-ray (with the *Einstein* observatory) and optical bands narrowed down the positional accuracy to $\sim 1''$ [283],[284]. But the final identification of Geminga with a pulsar came with *ROSAT* observations, which detected a periodic emission of 237 ms. Using the ephemeris data from *ROSAT*'s observations, the periodic signal was later identified in the EGRET as well as in the older *SAS 2* and *COS-B* data [285],[286],[287]. Astonishingly, the radio counterpart to the γ and X-ray pulses is very feeble, which makes high-energy emission this object's speciality. The γ-ray pulses from Geminga (> 70 MeV) show a double-peaked structure separated by two bridge components: one minor and one major; but the profile's characteristics change as one moves to lower energies [288].

Geminga's multi-wavelength spectrum is shown in Fig. 3.16, where one can see a turnover at ≈ 3 GeV, which is followed by a steep decline. Unfortunately, there has not been a detection of pulsed emission by VHE experiments so far, and thus we only have upper limits for the higher energies [289]. It seems, however, that the complexity of Geminga's pulse structure, together with the surprising fact that this pulsar's radio flux is, at most, much lower than the high-energy flux, makes this object a fair ground for the Outer Gap model, which places radio and high-energy emission at different locations in the magnetosphere [290].

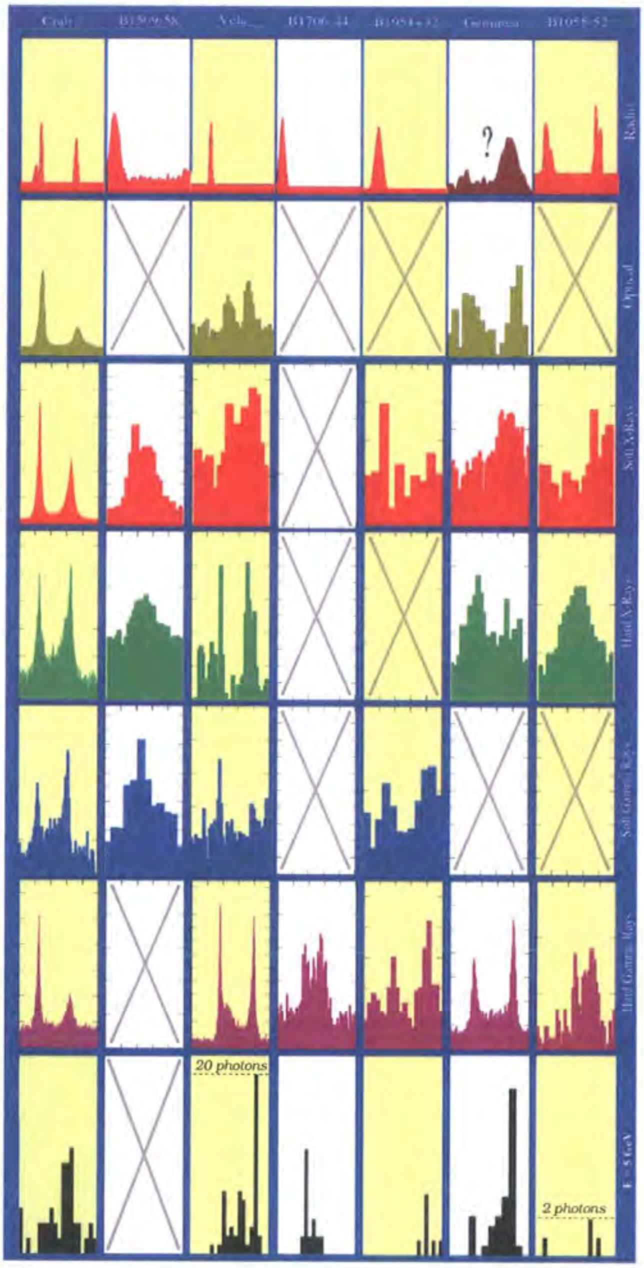


Figure 3.21: (from [269] and [279]) The pulse shapes of the 7 detected high-energy pulsars across the electromagnetic spectrum. The empty boxes correspond to non-detection at the respective energy range. Geminga's radio pulse has not been clearly detected and, so far, only claims exist (see e.g. [280]). PSR B1509−58 was only seen in X-rays up to 30 MeV, with COMPTEL, whereas the rest were detected above 100 MeV, with EGRET. The bottom row shows the pulse shapes at the top of EGRET's energy range (> 5 GeV).

PSR B1706−44 is a young, isolated, southern-sky pulsar in the constellation of Scorpio. It is of special interest for southern-hemisphere γ-ray observatories, like H.E.S.S. and CANGAROO, since it can be observed at small Z.A.s ($< 20°$) at culmination: a decisive factor for ground-based experiments, in terms of atmospheric attenuation of the Cherenkov light. Its period of rotation is 102 ms, but no VHE γ rays have been detected so far at this period. On the contrary, this pulsar has been prominent in γ rays below 20 GeV, and also in the radio band where it was first detected. Moreover, Gotthelf, Halpern and Dodson recently reported pulsed X-ray emission in *Chandra* data, but the detection was marginal ($\sim 4\ \sigma$), and the poor timing resolution of the high-resolution camera (HRC-I) on the *Chandra* satellite (4 ms) makes the detection even more uncertain [291].

So far the only high-confidence information we have for this pulsar's pulse shapes comes from the radio and γ-ray detections with EGRET. PSR B1706−44 produces a single bunch of emission in both cases. Nevertheless, as with almost all available γ-ray profiles, the pulse position is different for these two bands. Furthermore, the morphology is also different in that the γ-ray pulses are much broader and occupy $\approx 30\%$ of the period. A closer examination of the γ-ray pulse shows, however, that it is likely that the broad pulse seen above 100 MeV is a composite of two peaks separated by $\approx 70°$.

The pulsed spectrum of PSR B1706−44, as was derived from EGRET observations, is shown in Fig. 3.22, where one can notice that inside EGRET's energy range this pulsar does not appear to have a spectral turnover. This motivates the search for possible emission in the energy range of Cherenkov experiments. Both Polar Cap and Outer Gap models predict a sharp cut-off for this pulsar, at energies that are not very far above EGRET observations. Moreover, the Outer Gap model predicts an additional, much fainter TeV component, which is due to the inverse Compton emission from the gaps.

In addition to the direct study of the pulsar itself, PSR B1706−44 has also been the centre of a large debate regarding the existence of a possible surrounding nebula. The detection of unpulsed emission in both X and γ rays is in favour of such hypothesis [292],[34]. At present, the most likely candidate for a possible association is the nearby supernova remnant G343.1−2.3 [72],[293]. However there are arguments against this: see e.g. [294].

Vela (PSR B0833−45) is the brightest persistent γ-ray point source in the sky,

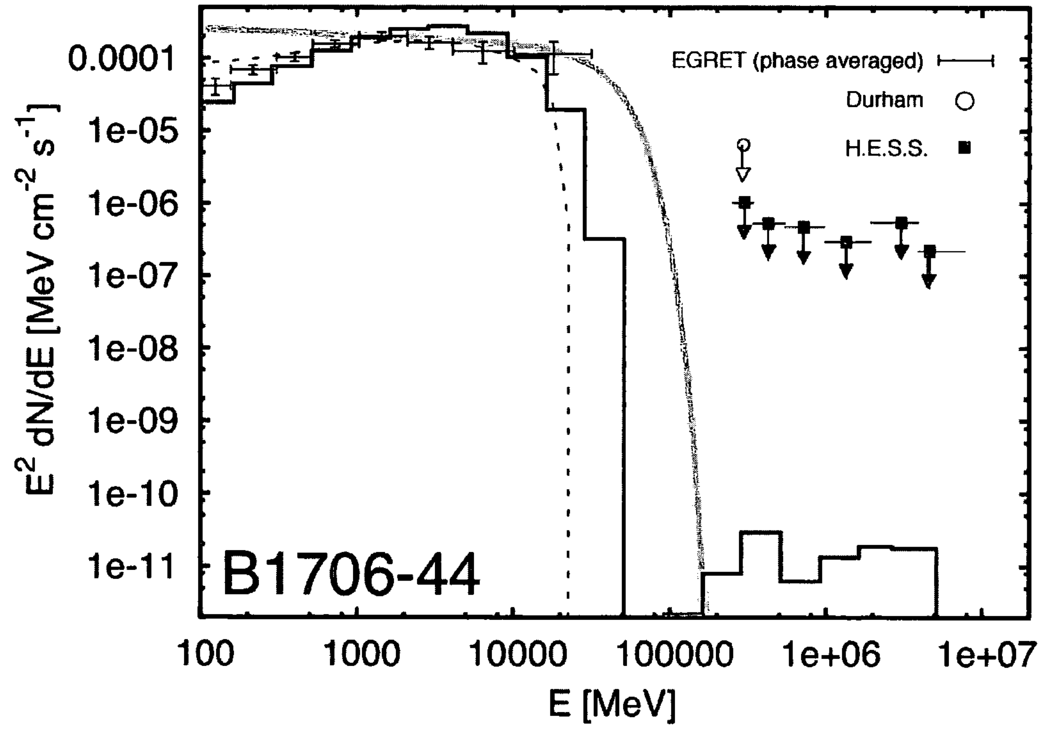


Figure 3.22: The phase-averaged spectrum for PSR B1706−44 from EGRET observations below 20 GeV. Also shown in this plot there are model spectra from the Outer Gap (solid line) and Polar Cap (dashed line) models, and a power-law fit to the EGRET data, with a super-exponential cut-off (thick grey line) [295],[296],[238]. The upper limits shown have been derived from observations with the Durham Mark 6 detector (circle) and H.E.S.S. (solid squares) [34],[256].

superceded only by the much brighter — but short-lived — GRBs. It is located in the Vela constellation, as its name suggests, and rotates with a period of 89 ms. It was born $\approx 12,600$ y ago in a supernova explosion whose remnants are still visible today in the optical band as a filamentary structure; and in X-rays, they appear as an expanding spherical shell.

The Vela pulsar appears very bright in radio and γ rays, and it is also detectable in optical and X-rays as a faint emitter. The widespread range of emission is also consistent with this pulsar's high spin-down power, which is second only to the Crab pulsar's (Vela and Crab share many features). In contrast to the early detection of this pulsar in hard γ rays with *SAS 2* and *COS-B*, hard X-ray and soft γ-ray emission (< 30 MeV) has only been recently reported with COMPTEL on-board *CGRO* [297],[277],[298].

Vela's pulsed γ-ray profile has a distinct double-peaked feature that is separated by 150° of bridge emission. Neither peak is coincident with the radio one, but both are as narrow. Moreover, the strength of the γ-ray signal for this pulsar allowed for spectral analysis of each component, separately. It was found that all three phase regions, which correspond to the two peaks and the bridge component, have different spectral characteristics, with the trailing peak having significantly softer spectral index than the rest. This feature is also consistent with the diminishing of the trailing peak with increasing energy, which is observed up to energies where data is still available (see Fig. 3.21).

The spectrum of Vela shows a sharp cut-off above 4 GeV, with which both theoretical models are consistent. However, although the Polar Cap models predict that the emission has a sharp cut-off above this energy, the Outer Gap models allow a slightly gentler fall of the spectrum, as well as an additional TeV component (see Fig. 3.23).

Another interesting feature of the spectrum derived from *ROSAT* observations is the soft X-ray component at ~ 1 keV energies. This is a spectral signature present in the spectra of 5 out of 7 EGRET pulsars and is thought to originate from the polar caps (see Fig. 3.16) [299].

The Crab (PSR B0531+21) pulsar is the energy source of the Crab nebula in the constellation of Cancer, and a strong source of emission in a broad range of frequencies, which spans from radio waves to γ-rays. Its rotation period is ≈ 33 ms, and its birth in 1054 AD — one of the few on record — makes this pulsar the

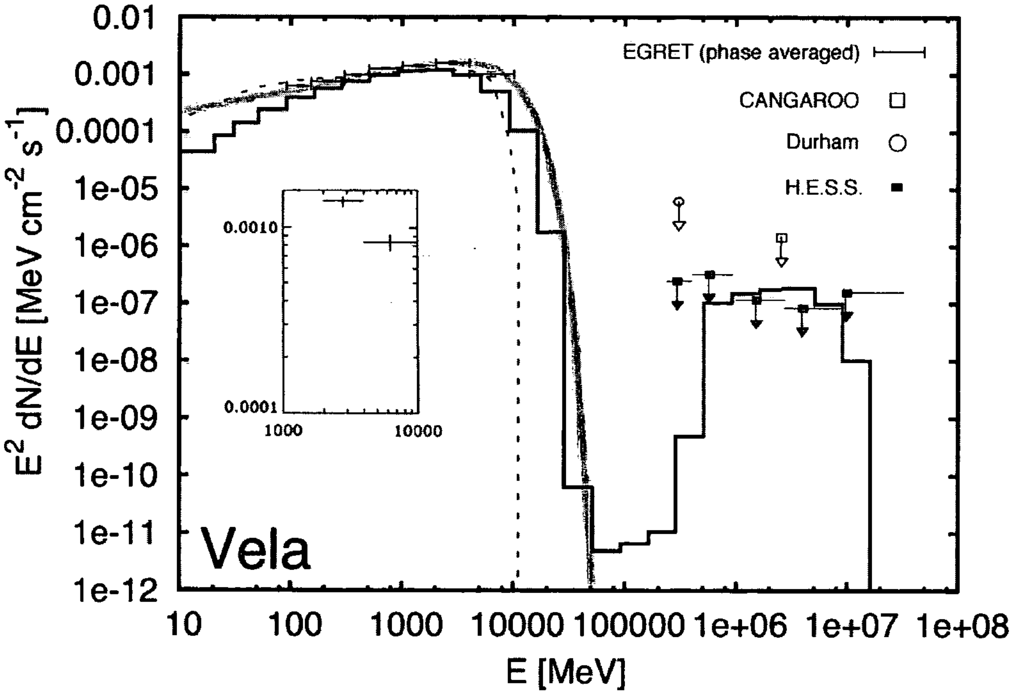


Figure 3.23: The phase-averaged spectrum of the Vela pulsar from EGRET observations below 10 GeV. Also shown in this plot, there are model spectra from the Outer Gap (solid line) and Polar Cap (dashed line) models, and a power-law fit to the EGRET data, with a super-exponential cut-off (thick grey line) [295],[296],[238]. The inset box shows the magnified area of the two top energy bins of EGRET. Upper limits on the pulsed flux have also been included [300],[71].

youngest γ-ray pulsar yet.

The Crab SNR boasts an impressive bolometric luminosity of $\sim 10^{38}$ erg s^{-1}. Yet, Crab pulsar's emission, like all the rest of the known γ-ray pulsars, stops somewhere in the VHE range, which is something not true for the surrounding nebula. The latter efficiently converts the abundant hadrons into TeV γ rays via pionisation. The detection of such unpulsed γ-ray emission has been achieved by virtually all modern VHE and UHE experiments: see e.g. [301],[302] and [303]. In addition, virtually all efforts to detect pulsed emission up to 20 GeV were successful, as opposed to very few reports of detectable, pulsed VHE γ rays.

The Crab pulsar's profile is unique amongst the EGRET pulsars, in the sense that it remains practically the same throughout the energy spectrum. Apart from the appearance of a precursor in the radio frequencies below 600 MHz, the phase of emission remains constant up to EGRET's energy range. At VHE above 5 GeV, as Fig. 3.21 shows, one can see the typical swap between the strengths of the first and second peaks. Although the latter phenomenon has not been explained yet, it seems that the phase consistency favours the Polar Cap models, in that both high-energy and radio emission originates from the same region. Nevertheless, there have been various arguments that support the coherent generation of radio and high-energy photons from the same $e^- - e^+$ pairs in Outer Gap models, too; although the explanation is more complex [304],[247].

The Crab pulsar's multi-wavelength spectrum appears continuous and relatively flat from the optical to γ rays and shows no evidence for the typical X-ray component that is present in most other high-energy pulsars. The EGRET data for the Crab pulsar can be well fitted with a single power law up to $\approx$ 10 GeV. Above this energy, only upper limits exist and the spectral shape is a subject for speculation. For example, based on a Polar Cap scenario where γ rays are attenuated by the strong magnetic field, one can assume a super-exponential cut-off at $\approx$ 30 GeV (see section 3.1.3). On the other hand, Polar Cap predictions become inconsistent at EGRET's top energies ($>$ 10 GeV), whereas the Outer Gap model is more consistent with the data and predicts, in addition, the typical for the model TeV component, which is, however, very weak in this case (see Fig. 3.15).

PSR B1055−52 is the oldest amongst EGRET γ-ray pulsars, with a spin-down age equal to $\log(\tau/y) = 5.7$. Its period is $\approx$ 197 ms, and its distance has been estimated to be $\approx$ 1.5 kpc, although there have been reports of lower values

[305],[306],[307]. This pulsar would be considered as the exception amongst γ-ray pulsars if the sample were representative of the population. Its observability, although high in absolute terms, places it below the top 28. Of course this pulsar *has* been detected, and therefore it is possible that the reason is its high γ-ray efficiency, or maybe that the distance to PSR B1055−52 is erroneous and that it is in fact much closer. However, if one trusts the current derived values, then this pulsar's efficiency is about 5 times higher than the EGRET average, with an impressive $\epsilon \approx 15\%$. Alternatively, another possibility is that the sites of radio and high-energy emission are such that the latter is favoured over the former, which is also something that one may consider true for another puzzling case: the Geminga pulsar.

The radio profile of PSR B1055−52 has nothing in common with that at HE. In contrast to the radio pulse, γ-ray emission is broad and contained in a single, roughly contiguous phase range, although a double-peaked morphology can be distinguished. Together with PSR B1706−44, PSR B1055−52 exhibits the broadest pulse amongst the EGRET pulsars. Along the lines of a Polar Cap scenario, these kind of features can be explained by the different emission regions in the hollow cone model, whereas the Outer Gap model accounts for them mainly in terms of the difference in the time-of-flight between beams produced at different locations in the gaps.

Finally, the spectrum of PSR B1055−52, as was measured by EGRET, does not show a steep decline within the energy range of the observations, but the upper limits from VHE experiments, like CANGAROO, imply a cut-off in the spectrum somewhere in between. In addition, like all the rest of the γ-ray pulsars, PSR B1055−52's luminosity peaks in γ rays. But contrary to the rest, its efficiency is much reduced in other wavelengths, which is a characteristic that is believed to be typical for old pulsars (see Fig. 3.16).

PSR B1951+32 is approximately 10,000 y old and has the weakest magnetic field of all discovered γ-ray pulsars, with $log(B_s/G) \sim 11.7$. It is also the most distant γ-ray pulsar EGRET has discovered, which is part of the reason for the very low γ-ray flux received from this object. It was discovered in radio with a period of 39.5 ms, inside the synchrotron nebula CTB 80 [308]. However, its detection in the MeV–GeV range was only accomplished some ten years later with EGRET [25].

The pulsed profile of PSR B1951+32 above 100 MeV shows a clear separation between two peaks, neither of which is coincident with the single radio pulse. Unlike the rest of EGRET pulsars, there is no statistically significant evidence for bridge emission anywhere within the ≈ 160° that separate the leading and trailing pulses.

PSR B1951+32 is a good candidate for VHE γ-ray observations because of its hard spectral cut-off, which, according to EGRET data, must be above $\approx (20 \pm 10)$ GeV. Northern-hemisphere observatories (e.g. VERITAS and MAGIC), in particular, will have the advantage of PSR B1951+32 observations at small Z.A.s, which will better their chances of detection with regards to pulsed emission beyond EGRET energies.

PSR B1509−58 is not an EGRET pulsar, but it has been detected with COMPTEL up to 30 MeV. It was first discovered in X-rays with the *Einstein* observatory during the study of the supernova remnant MSH 15−52, with which it was later associated after radio observations. Its period of rotation is 150 ms, and its characteristic age, $\log(\tau/y) = 3.2$, makes this pulsar the second youngest amongst EGRET's. Although this pulsar's position makes it a good target for southern observatories, the fact that there have only been upper limits in the EGRET range probably implies no emission at VHE energies, unless there exists a detectable inverse Compton component at TeV energies [309].

The pulse of PSR B1509−58 has a single, broad profile in the 0.75–10 MeV energy range, but at higher energies (up to 30 MeV) this profile becomes substantially weaker. However, there exists a very narrow, significant component at phase 0.85, which only appears above 10 MeV. Beyond COMPTEL's range, there is no significant event contribution at the expected pulse position, although the contribution from the single, narrow pulse at phase 0.85 still persists and causes a spectral steepening above 30 MeV (see Fig. 3.24).

Fig. 3.25 shows a plot of the high-energy cut-offs of the 7 γ-ray pulsars against their surface magnetic fields. PSR B1509−58 has the lowest cut-off, amongst the discovered sample, and the highest surface magnetic field. In the grounds of a Polar Cap framework, this relationship is well explained by the increased efficiency of magnetic pair production in strong magnetic fields ($\gamma + \boldsymbol{B} \longrightarrow e^- + e^+$), which leads to lower escape probabilities for energetic γ rays. Similarly, the rest of the pulsars, having weaker magnetic fields, should be expected to have higher cut-offs. Although this appears to be true for the currently available sample, the low statistics and the undetermined cut-off values in more than half of the cases does not allow for any trend to be established across the range of magnetic fields. As more pulsars are discovered by future experiments, a statistical analysis may well show a correlation between magnetic-field strength and cut-off energy; but the current state of knowledge, as is presented in Fig. 3.25, could either be a hint to the Polar Cap's predictions, or it could also be misleading, with PSR B1509−58 being a statistical outlier.

3.2.3 Hidden Behind their Ashes

Plerion versus Pulsar Emission: A Collaborative Contrast

From a high-energy astrophysicist's perspective, pulsars and their surrounding environments are related in the most interesting way. Although their observable properties appear quite different, the underlying physical processes reveal a common nature. After all, both the central compact object and the expanding shell of gas were once parts of the same thing: the progenitor giant star.

There is clear evidence that supports an associated past between a pulsar and its surrounding nebula. For example, it was mentioned in section 3.1.3 that the energetics of plerions, like the Crab nebula, can be explained as being the result of the Crab pulsar's total energy output. In addition, all previously discussed γ-ray pulsars are coincident with a supernova remnant, except perhaps for the radio-quiet Geminga, for which there is yet unclear evidence for an extended X-ray source surrounding the pulsar [315],[316],[317]. In cases like that of PSR B1951+32, observation of the plerion resulted in its association with a serendipitous detection of an embedded pulsar, and vice versa (e.g. in the case of PSR B1706−44; see section 3.2.2). Plerions and pulsars are tightly bound by their history.

The Crab and Vela pulsars are examples of such plerionic pulsars: their respective nebulae interact with the pulsar wind and give rise to synchrotron emission which emanates from the confined pulsar wind, within the limits of the nebula. Although direct pulsed emission may not be visible because of the geometrical configuration with respect to the Earth, the nearly omnidirectional plerionic emission can be detected in a broad range of frequencies, from radio to γ rays. For energies up to X-rays, the relativistic electrons can account for the observed synchrotron spectrum, but the VHE regime is chiefly due to inverse Compton and pion decay components (see sections 2.5 and 2.6). The former component comes from within the nebula's boundaries, from the collisions between IR background and synchrotron-emitting electrons, whereas pion decay occurs outside the nebula, from the collisions of shock front–accelerated hadrons with static ISM hydrogen.

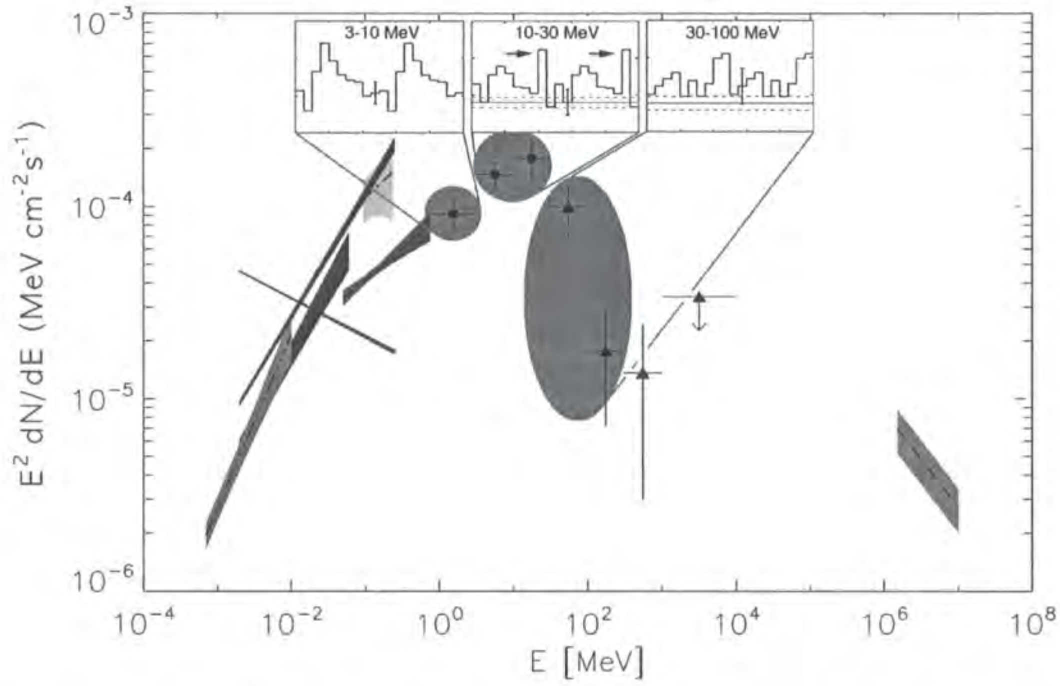


Figure 3.24: (from [310]) The high-energy spectrum of PSR B1509−58 from COMPTEL and EGRET observations. The shaded, circular areas define three energy ranges in which the COMPTEL (solid rectangles) and EGRET data (solid triangles) were split in order to investigate the pulse modulation with energy. One can see that above 10 MeV (in the COMPTEL range) the event contribution at the expected phase range disappears, and only the contribution from a narrow component, at ϕ =0.85, is present (black arrow). The spectrum of this narrow component is different to the pulses at lower energies, which causes an increase in the flux levels. Above 100 MeV (in the EGRET range), both broad and narrow emission disappears, and the spectrum is consistent with the unpulsed, nebular γ-ray spectrum that was detected at TeV energies with CANGAROO (shaded area above 1 TeV) and *RXTE* (shaded areas below 1 MeV).

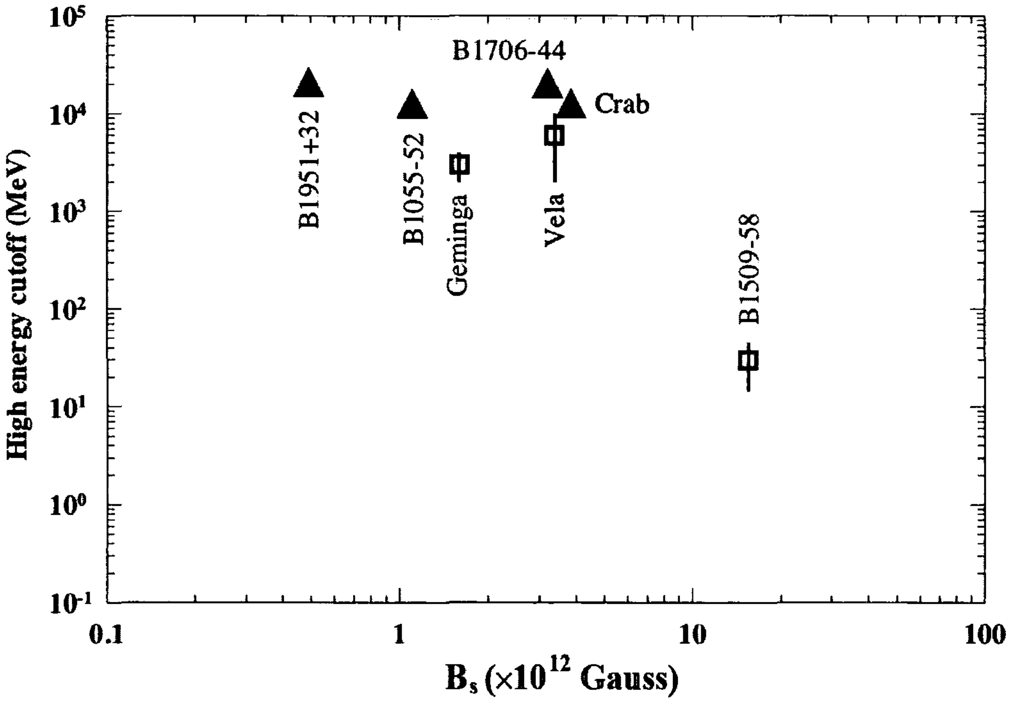


Figure 3.25: (from Table 3.3) Plot of the high-energy cut-offs of EGRET pulsars against their surface magnetic fields, B_S. PSR B1509−58, which was detected with COMPTEL up to 30 ± 15 MeV, is also shown. The solid triangles correspond to lower limits, as there is little evidence for a turn-over in the corresponding pulsar spectra, within EGRET's energy range.

	Geminga	B1706−44	Vela	Crab	B1055−52	B1951+32	B1509−58
P (ms)	237	102	89	33	197	39	150
$\dot{P}$ (10^{-15} s s^{-1})	11.0	93	125	421	5.83	5.85	1,600
[1]ϵ_γ (%)	5 ± 4	0.72 ± 0.47	0.18 ± 0.07	0.013 ± 0.006	15 ± 7	0.26 ± 0.17	[a]0.9
$\log B_s$ (G)	12.2	12.5	12.5	12.6	12.0	11.7	13.2
Spectral cut-off (GeV)	3 ± 1	$\gtrsim 20$	6 ± 4	$\gtrsim 10$	$\gtrsim 10$	$\gtrsim 20$	0.03
[2]$\dot{E}_{\mathrm{rot}}$ ($\times 10^{36}$ erg s^{-1})	0.033	3.4	7	450	0.03	3.7	18
$-\log(\dot{E}/d^2)$ (erg cm^{-2} s^{-1})	7.3	7.0	5.5	4.9	8.9	7.0	
[3]γ	1.42	2.1	1.62	2.08	1.8	1.74	-
[4]$\log \tau$ (y)	5.5	4.2	4.1	3.1	5.7	5.0	3.2
Distance (kpc)	0.2	1.8	0.5	2.0	1.5	2.5	4.4
[5]F_{ul} (10^{-12} cm^{-2} s^{-1})	[6]6.54	[7]8	[8]0.37	[9]4.82	[10]68	[11]6.7	[12]1.5
Discovery reference	[26]	[24]	[23]	[22]	[5]	[25]	[310]

Table 3.3: The properties of the 7 high-energy pulsars and the corresponding VHE upper limits on pulsed emission from various campaigns.

[1] The γ-ray conversion efficiency is the γ-ray fraction (> 100 MeV) of the spin-down luminosity [228]. [a]For PSR B1509−58 the high-energy conversion efficiency has been calculated for all energies above 1 eV, i.e. above the optical band.

[2] Spin-down luminosity, $\dot{E} = 4\pi^2 I\dot{P}/P^3$ [5].

[3] The derived differential spectral index in the EGRET energy range.

[4] The logarithm of the dynamic age $(1/2)(P/\dot{P})$.

[5] The VHE integral flux upper limits on pulsed emission.

[6] Whipple upper limit for $E > 500$ GeV at 5-σ confidence level [311].

[7] CANGAROO flux of the total emission (pulse and unpulsed) above 1 TeV [312].

[8] CANGAROO upper limit for $E > 2.5$ TeV at 95% confidence level [71].

[9] Whipple upper limit for $E > 250$ GeV at 99.9% confidence level [95].

[10] Durham Mark 6 upper limit for $E > 300$ GeV at 3-σ confidence level [313].

[11] Whipple upper limit for $E > 260$ GeV at 99.9% confidence level [314].

[12] CANGAROO upper limit for $E > 2.5$ TeV at 3-σ confidence level [68]

Hence, the observed γ-ray emission from plerions has properties that are clearly distinguishable from the central compact object's. The direct emission from pulsars occurs somewhere above the polar caps, either close to the star (Polar Cap) or on the edge of the light cylinder (Outer Gap), and therefore follows the neutron star's rotation and appears to a static observer as pulsed. On the other hand, the plerionic emission does not exhibit a preferred direction and shows no flux variability within the experiments' limitations (see e.g. [318]). Furthermore, the flux of the pulsed emission for most pulsars peaks somewhere inside the EGRET energy range and, evidently, shows a steep decline at energies of a few GeV, whereas the plerionic emission is much weaker at those energies and is persistent up to $\sim$ 100 TeV.

Fig. 3.26 shows Crab's pulsed and DC spectra as they were derived from observations with EGRET and ground-based VHE experiments. One can see the clear distinction between the low-energy synchrotron and VHE IC components of the plerionic emission (also see Fig. 2.5). Also notable is the prominence of the pulsed emission in the energy range from 100 MeV to 10 GeV, whereas the plerionic DC emission becomes important at higher energies. Once again, it is made clear that by pushing the sensitivities of ground-based experiments towards the cut-off energies, one can hope to detect the pulsed synchrotron flux, which is thought to show a steep (e.g. exponential) increase somewhere below 100 GeV.

Most experiments have already detected Crab's nebular DC emission. Its IC component is clearly detectable above the threshold of modern detectors (> 100 GeV). Imaging and stereoscopy conspire towards a better γ/hadron discrimination and, thus, help discard a large fraction of the background (i.e. cosmic rays, muons, etc.). For such observations, the signal strength is measured by counting the surviving on-source (ON) and off-source (OFF) events and calculating the quantity

$$\sigma_{\mathrm{DC}} = \frac{\mathrm{ON}-\mathrm{OFF}}{\sqrt{\mathrm{ON}+\mathrm{OFF}}} = \frac{R_\gamma}{\sqrt{R_\gamma + 2R_\mathrm{c}}} \cdot \sqrt{T} \tag{3.29}$$

which gives the significance (as a number of standard deviations) of the DC γ-ray signal, assuming a Poissonian background. In Eq. 3.29, R_γ and R_c are the γ-ray and background event rates, respectively. Note also that for a constant γ-ray rate, R_γ, the signal strength increases as $\sigma_{\mathrm{DC}} \propto T^{1/2}$, where T is the exposure time. Finally, the factor 2 in front of the background rate is due to the involvement of the background events in both ON and OFF observations.

On the contrary, none of the ground-based experiments have produced a solid detection of pulsed emission. If the pulsed emission models are correct, the present detectors will only be capable of a marginal detection of such signals during their operation periods: this is because the energies where one has to look for periodic emission are expected to be well below the energy thresholds of the currently operating experiments ($\lesssim$ 50 GeV). Firstly, at those energies, the detection efficiency is rapidly reduced, and therefore the ability to retain a statistically significant fraction of the occurring showers — usually quantified in terms of effective area — is low (see Fig. 3.26). Secondly, the background rejection is poor because the imaging technique is not applicable for such faint showers, and hence one has to include whatever background events are detectable in this part of the spectrum, in order to not risk losing valuable pulsed photons which have not been recognised as γ rays.

Fortunately, the temporal character of the periodic emission can be used to our advantage: the available information from lower energy observations on pulsar properties (i.e. the period, pulse position, etc.) helps us to search for a periodic signal at the expected frequency, or even at the specific phase where we know the pulses are likely to appear. In this way, the statistical significance of a potentially existing signal is enhanced over the uniform background. The calculation for this significance is not as straightforward as that for a DC analysis, since it involves a few assumptions about the position and shape of the pulses. More about periodicity searches with different tools is discussed in section 5.4.

In this paragraph, we discussed, in brief, the physical relation between pulsars and their surrounding nebulae but also addressed the fundamental differences in the observed emission from both. Although the data analysis for either object should follow a different strategy, it makes sense to try and estimate which percentage of the pulsar's radiation is re-emitted at higher energies from the nebula. For this purpose, a useful quantity is the ratio ${F_\gamma}^{\mathrm{pulsed}}/{F_\gamma}^{\mathrm{DC}}$, which gives the pulsed γ-ray flux relative to the DC emission. However, as is more common, information on either fluxes might not be available: e.g. in the case of non-detection of either the pulsed or DC component, or even when the energy range where the pulsed emission is detectable is below the telescope's threshold for a clear DC detection. In those cases, the common approach is to estimate an upper limit on the unknown component, F_{ul}, and calculate the ratio

$$f\% = \frac{F_{\mathrm{ul}}}{F_{\mathrm{known}}} \cdot 100\% \tag{3.30}$$

where F_{known} is the flux of the known component. Unfortunately, sometimes even

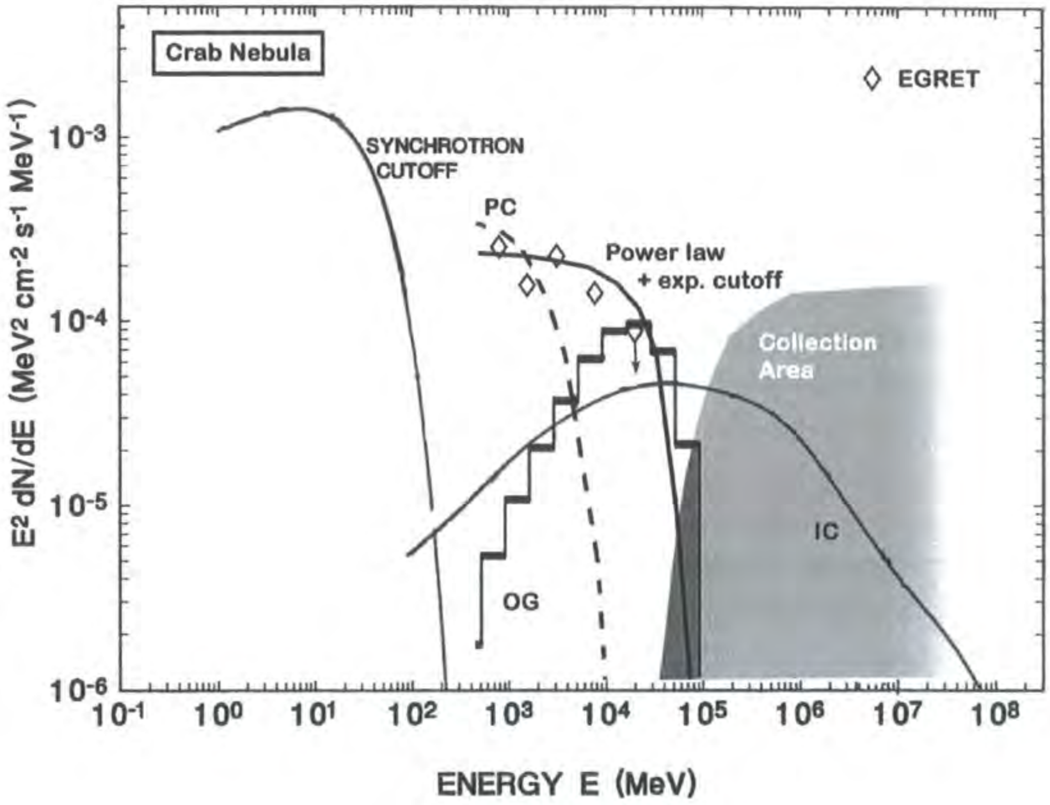


Figure 3.26: The Crab nebula's synchrotron and IC components. Also shown in this plot are two model spectra for the Crab pulsar's emission, and an additional power-law fit to the EGRET data, with an exponential cut-off at ≈ 10 GeV. For details on the specific results for the Crab nebula's spectrum, from the various experiments, see [28] and references therein. The lightly shaded area which extends to high energies represents a typical dependence of the collection area on energy, for contemporary Air Cherenkov experiments. It can be seen that by reducing the telescope's energy threshold, it becomes possible to detect the tail-end of pulsed emission. The more heavily shaded area represents the detectable portion of the pulsed emission, if one assumes Outer Gap emission with a typically slow roll-over.

that can be difficult to calculate!

Chapter 4

Very High Energy γ-ray Astronomy

4.1 The Earth's Atmosphere: A Gigantic Photomultiplier

4.1.1 Introduction to the Principles of Extensive Air Showers

The energy source of the Crab nebula — which lies at a distance $d \approx 6.3$ kly from the Earth — is a ~ 1.4-$M_{\odot}$ pulsar whose spin-down luminosity is $L_{\rm sd} \sim 10^{38}$ ergs $\rm s^{-1}$. Only a small percentage of this luminosity is responsible for the total γ-ray emission from the plerion. The amount of γ rays that we receive on Earth is $F_{\gamma}(> 100 \text{ MeV}) = 2.45 \times 10^{-6}$ $\rm cm^{-2}$ $\rm s^{-1}$, which leads to a kinetic-energy-to-γ-ray conversion factor of the order

$$\epsilon = \frac{4\pi f d^2 F_{\gamma} \langle E \rangle}{4\pi^2 I \dot{P} / P^3} \sim 10^{-4} \tag{4.1}$$

where the denominator is the spin-down luminosity of the Crab pulsar; the numerator is the γ-ray luminosity of the beamed emission, with beaming factor assumed equal to $f = 1/(4\pi)$; and $\langle E \rangle$ is the average photon energy over the observed range [318].

Nevertheless, it is the last part of the radiation's journey that is responsible for the visibility of VHE γ-rays from the ground. Each primary γ ray — or simply *primary* — that enters the Earth's upper atmosphere with an energy in the GeV–TeV range pair-produces after one radiation length, X_0. This corresponds to the distance in the traversed medium, across which a particle loses all but 1/e of its initial energy via Bremsstrahlung. This characteristic length is roughly the same for pair production and Bremsstrahlung and, in the atmosphere, is equal to $X_0 = 37.1$ g cm^{-2}, which corresponds to an altitude of ≈ 20 km [1].

After pair production has occurred, the primary converts its energy to kinetic energy of an $e^- - e^+$ pair. The pair travels nearly parallel to the primary's direction at relativistic velocities. As it makes its way towards the lower atmosphere, it produces further amounts of γ rays via Bremsstrahlung. These secondary γ rays can pair-produce again, provided they are energetic enough. The photon–particle cascade that develops this way in the atmosphere is called an extensive air shower (hereafter EAS), or just air shower. In particular, cascades whose products are photons, e^- and e^+ are called EM cascades.

Fig. 4.1 shows a simplified model of an EM cascade, where the primary γ ray is converted to a ramification of particles and secondary γ rays. The cascade products interact with the atmosphere in discrete distance intervals that are equivalent to X_0. Under the assumption that the products from each interaction share the energy of the parent particle or photon equally amongst them, one can express their energy at every step of the cascade with

$$E_n = E_0 2^{-n} \tag{4.2}$$

where E_0 is the energy of the primary, and n is the integer number of radiation lengths between the primary and the n^{th}-in-a-row interaction — when each product's energy is E_n. By invoking the radiation length, the *average* particle energy as a function of the distance, $x = nX_0$, traversed in the atmosphere becomes

$$\langle E \rangle = E_0 2^{-x/X_0} \tag{4.3}$$

To be more specific, the quantity x expresses the amount of matter traversed in the atmosphere in units of g cm^{-2} and is called atmospheric depth.

The succession of light-to-matter and vice versa conversions reaches its maximum when the energy losses due to Bremsstrahlung radiation become equal to ionisation losses due to the collision of electrons with atmospheric nuclei. The critical energy at which this occurs in the atmosphere is $E_{\text{crit}} \approx 80$ MeV. At that point, the EM shower reaches the so-called *shower maximum*. Eventually, the shower dies out when the average energy of the produced photons drops below the threshold for further pair

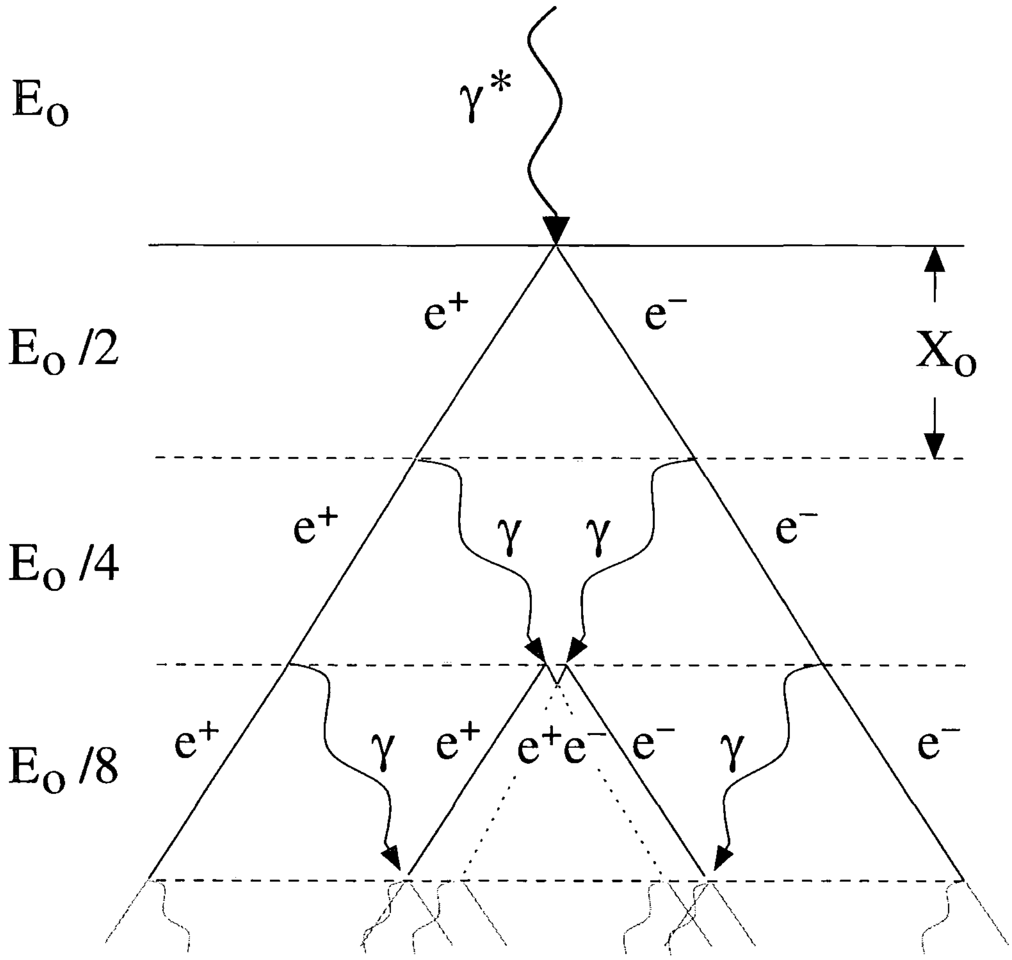


Figure 4.1: A simple model of EAS development. Each particle or photon of the cascade interacts with the atmosphere via Bremsstrahlung or pair production, after traversing one radiation length, X_0. The energy, E_0, of the primary γ ray (γ^*) is divided in equal shares amongst the successive products of the cascade (e^-, e^+, γ).

production. In the VHE range, such showers develop along the projected axis of the primary's direction, and the products of the cascade are tightly bunched around this axis. The typical longitudinal extent of 1-TeV air showers ranges from an altitude of $\approx$ 20 km to $\approx$ 4 km a.s.l. The shower maximum for VHE EAS can be anywhere between 10 and 6 km a.s.l., where the specific altitude depends on the primary γ-ray energy. For 1-TeV primaries, for example, this occurs at 8.4 km [1]. Fig. 4.2 shows a simplified schematic of an EM cascade's development from a 1-TeV primary.

As the shower particles traverse the atmosphere, their charge polarises the atoms of the medium and, hence, makes them electric dipoles with dipole vectors pointing towards or away from the charged particle, depending on the sign of the charge. In the case where the charged particle has low energy, which is true for traversing particles created lower in the atmosphere, the atmospheric atoms have enough time to adjust to the potential field of the particle, and so there is a spherically symmetric arrangement of dipoles around the low-energy particle. On the other hand, the e^-–e^+ pairs produced higher in the atmosphere travel at nearly the speed of light, c, which causes the atmospheric atoms to respond to the charge particle's retarded potential, which is asymmetric around the true position of the particle. The form of this potential, also known as the scalar Liénard–Wiechert potential [185], is

$$\phi(\boldsymbol{R}_\mathrm{n}, t_0) = \frac{1}{4\pi\epsilon_0}\frac{e}{(R_\mathrm{p} - \boldsymbol{R}_\mathrm{p}\cdot\boldsymbol{v}/c)} \tag{4.4}$$

where $\boldsymbol{R}_\mathrm{n}$ is the position vector of one of the surrounding atoms where the potential is $\phi(\boldsymbol{R}_\mathrm{n}, t_0)$; and $\boldsymbol{R}_\mathrm{p}$ is the position vector of the particle causing the EM disturbance, calculated at the retarded time $\tau = t - |\boldsymbol{R}_\mathrm{n} - \boldsymbol{R}_\mathrm{p}|/c$ (see Fig. 4.3).

After the low- or high-energy particle has passed the part of the atmosphere where the disturbance was caused, the atoms return to their natural states. A minor EM wave is then emitted from every depolarised region along the particle's path. For particles that travel slower than the speed of light in the medium, the generated spherical waves propagate through the medium without ever intersecting (see inset of Fig. 4.4). However, superluminal charged particles cause the generation of EM waves, whose wave-fronts interfere constructively towards an enhanced emission that propagates in the forward direction (Fig. 4.4). That EM emission is called Cherenkov radiation and it is always visible whenever a particle, capable of polarising the atoms of a medium, is travelling faster than the speed of light in that medium: i.e. when $v_e > c/n$, where n is the refractive index of the medium.

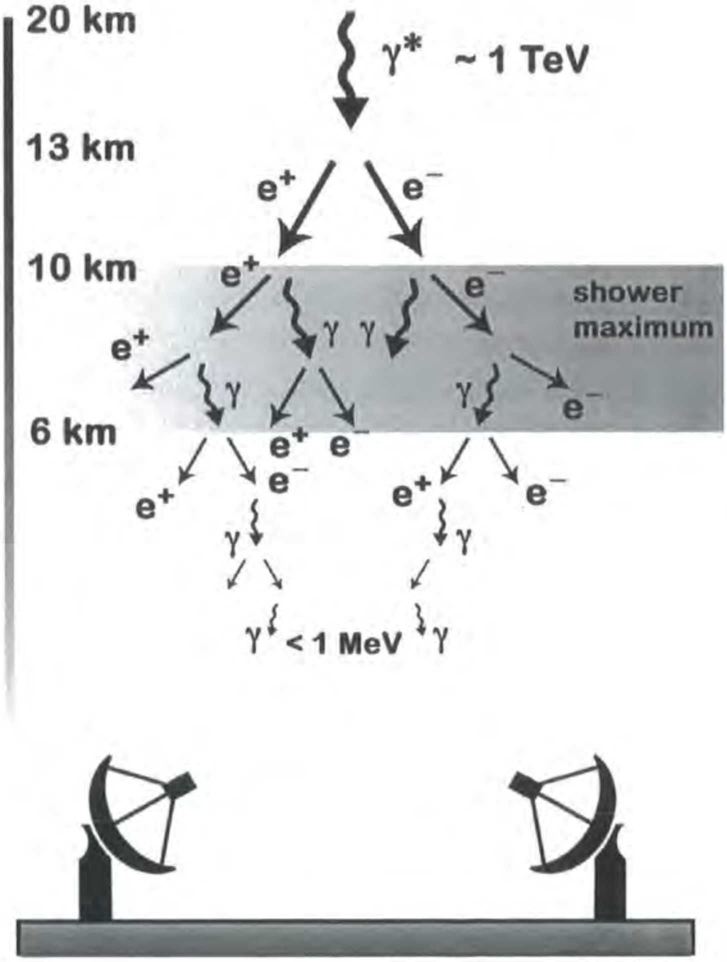


Figure 4.2: A 1-TeV, γ-ray-initiated EAS. The primary photon pair-produces soon after traversing one radiation length in the atmosphere (at ~ 20 km a.s.l.), and the products give off secondary γ-ray photons via Bremsstrahlung. These photons can pair produce again as the cascade develops. The maximum production of particles occurs somewhere between 10 and 6 km altitude. Finally, the cascade dies off as the subsequent absorption and re-emission processes leave out only thermal particles that can radiate no further. Please note that the horizontal scale has been enlarged by many times for clarity.

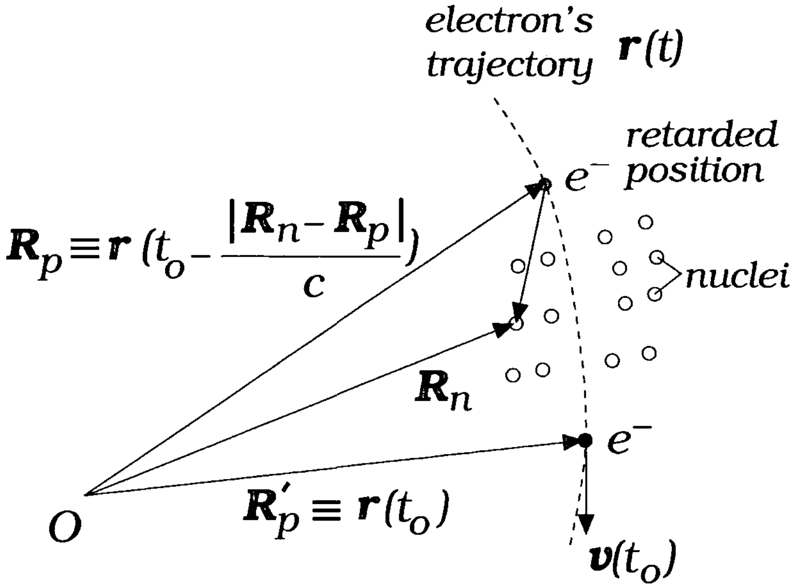


Figure 4.3: The definitions of the vectors involved in the calculation of the Liénard–Wiechert scalar potential. The trajectory of the electron is defined by the position vector, $\boldsymbol{r}(t)$, which connects the observer (O) with the electron's position. At time t_0 this vector is $\boldsymbol{R}'_{\rm p}$, whereas at an earlier time $t_0 - |\boldsymbol{R}_{\rm n} - \boldsymbol{R}_{\rm p}|/c$ (where $\boldsymbol{R}_{\rm n}$ is the position vector of one of the surrounding atoms) the electron's position vector is $\boldsymbol{R}_{\rm p}$. The time interval between the two positions is equal to the time it takes for the EM disturbance to reach the position of the atom where the potential is equal to $\phi(\boldsymbol{R}_{\rm n}, t_0)$.

In reality, Cherenkov radiation is emitted at an angle to the particle's velocity vector. Using simple geometry, one can show that the faster the particle — travelling in the medium — the wider the angle of emission; and that the greater the refractive index of the medium, the wider the emission angle from a particle travelling at constant velocity. One can derive these dependencies based on the drawing in Fig. 4.4.

At $t = 0$, we assume that the charged particle disturbs the surrounding molecules and causes the generation of the first spherical Cherenkov wave. After time t, the particle travelling at nearly the speed of light in the vacuum, βc, has covered distance equal to βct. At the same time, the first generated spherical Cherenkov wave has propagated, at the speed of light in the medium, to a radius equal to $(c/n)t$. In the case of a high-energy particle, we have that $\beta > 1/n$; and therefore the spherical waves cannot catch up with the particle — which is of course the main reason why there is Cherenkov radiation produced in the first place — and thus leave behind a trail of intersecting spherical waves which constructively add up to a plane wave with $\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{\beta}} = \cos\theta$. In this equation, $\hat{\boldsymbol{k}}$ is the unit vector of the wave's direction of propagation, and $\hat{\boldsymbol{\beta}} = \hat{v}_e/c$ is the particle-velocity unit vector normalised to the speed of light in vacuum. In 3-dimensional space, the Cherenkov emission is conical because of the axial symmetry around the particle's trajectory. The angle at which the Cherenkov photons are emitted with respect to the particle's velocity is

$$\theta = \cos^{-1}\left(\frac{ct/n}{\beta ct}\right) = \cos^{-1}\left(\frac{1}{\beta n}\right) \tag{4.5}$$

where n is the refractive index of the medium.

As an EAS develops in the atmosphere, the average energy of the shower's particles decreases due to radiation and ionisation losses [1]. The speed of these particles, βc, decreases with traversed distance until $\beta_{\text{min}} = 1/n$, at which point Cherenkov emission ceases. Hence, Eq. 4.5 implies that for a developing EAS in a medium with constant n (e.g. EAS that graze the atmosphere), θ tends to get smaller and smaller until, of course, $\theta_{\text{min}} = 0°$. In practice, a detected EAS always traverses the different density layers in the atmosphere, which means that the refractive index gradually changes from $n_{\text{min}} \approx 1$ to $n_{\text{max}} \approx 1.000293$ (near sea-level). In those cases, observations show that air showers tend to spread out (θ broadens) as they develop towards lower altitudes, which means that, during EAS development, the

---

[1] Cherenkov emission dissipates energy as well, but is only responsible for less than 1% of the total loss [319].

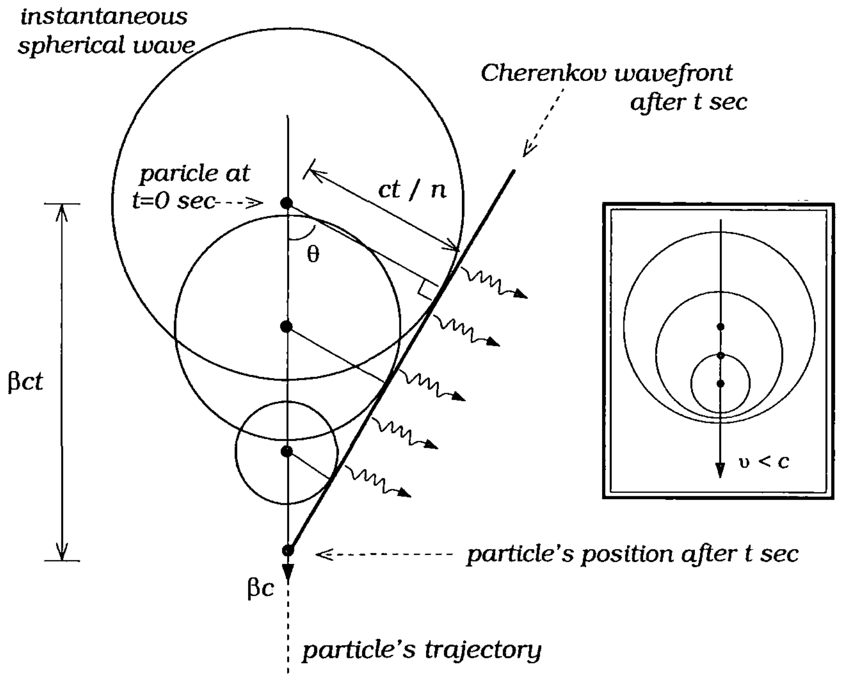


Figure 4.4: The Cherenkov emission geometry. As the charged particle propagates into the medium at superluminal speeds, it continuously generates spherical Cherenkov waves whose surfaces intersect. The constructive interference of all the successive waves created along the particle's path generates a plane wave of Cherenkov emission (thick, black line) which propagates at an angle $\theta = \cos^{-1}[1/(\beta n)]$ with respect to the particles velocity. In a 3-dimensional representation, this wave would appear as conical due to the axial symmetry of the phenomenon. The inset on the right of this figure shows the opposite case of a particle travelling slower than light. The generated waves propagate in the medium without interfering, and no detectable emission is observed.

rate of change of the refractive index is higher than the energy loss of the particles. Finally, at very low altitudes, where the last production of Cherenkov light is still possible, the produced light by the few remaining particles does not have time to spread substantially. As a result, the light from those local showers falls very close to the shower axis.

Fig. 4.5 illustrates the average paths followed by the Cherenkov photons and shower particles on their journey to the ground. At the top of this sketch, the 1-TeV primary interacts for the first time at 20 km. The dotted, vertical line depicts the mean path of the traversing $e^{-}-e^{+}$; the solid, slanted lines, the path of the Cherenkov photons travelling from the production region to the ground. The main things to be noted are the difference between the emission angles at different altitudes, which is due to the dependence of the refractive index on altitude, and the dotted box that envelopes the shower maximum, where more than 50% of the shower's total light is produced. The majority of the recorded EAS are detectable because of the emission from this region. It is worth mentioning that because of the different refractive indices between the upper and lower boundaries of the shower maximum, light produced on the lower boundary will be more spread out than that from the highest boundary. As a result, the geometrical paths of Cherenkov light produced inside the shower maximum converge (see Fig. 4.5). Consequently, light which arrives at the ground from this region is focused onto a thin annulus. Here, we provide an example calculation that gives an estimate of the Cherenkov light pool's extent at sea-level:

> Our example considers a 1-TeV EAS whose shower maximum occurs at ≈ 8.4 km a.s.l. Using Eq. 4.5, we will calculate the angle of Cherenkov emission from a single electron at the shower maximum. The average particle energy in this region is roughly 80 MeV, which corresponds to the energy threshold for ionisation in the atmosphere. Hence, the relativistic β for the electron is
>
> $$\beta \approx 1 - \frac{1}{2\gamma^2} \approx 1 - \frac{1}{2}\left(\frac{m_e c^2}{E_e}\right)^2 \tag{4.6}$$
>
> where $m_e c^2 = 511$ keV is the electron rest energy; and E_e, the relativistic electron energy. Setting $E_e = 80$ MeV, we get $\beta \approx 0.99998$.

The calculation of the refractive index at the position of the shower maximum requires an assumption about the atmosphere's density profile. Here, we

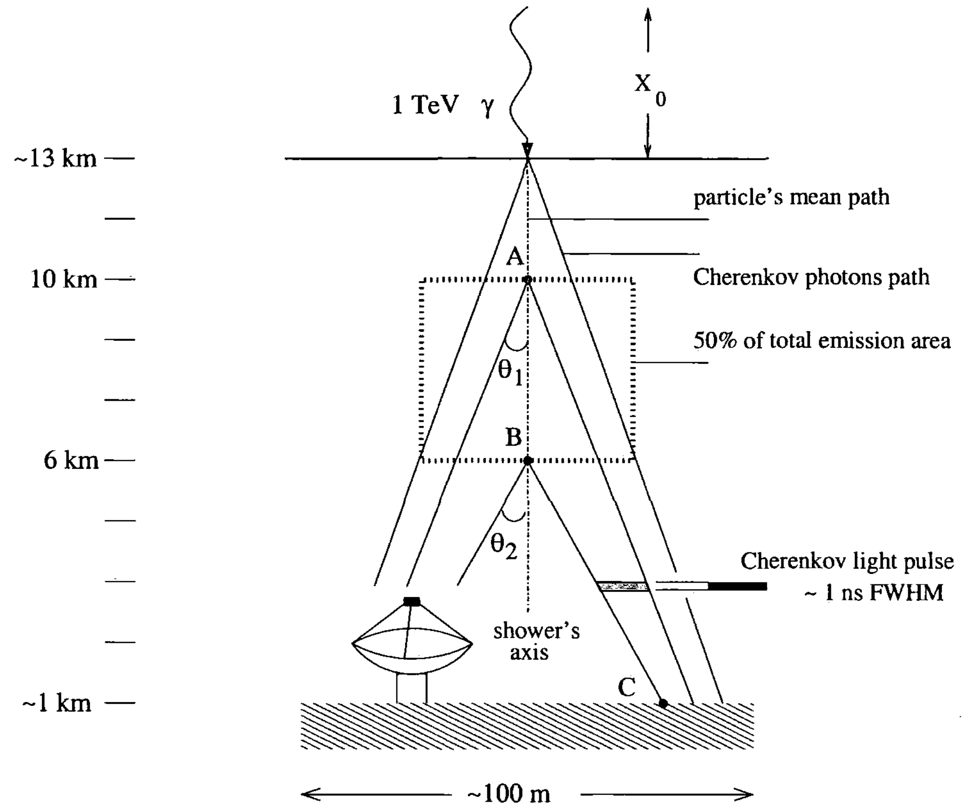


Figure 4.5: The focusing effect of atmospheric Cherenkov light caused by 1-TeV primaries. 50% of the shower's Cherenkov light is produced inside the area that is approximately defined by the dotted rectangle and is also known as the shower maximum. Cherenkov emission from the top of the shower maximum tends to be broader than that produced at the bottom region, which causes the top, bottom and all light beams in between to converge. In addition, the superluminal velocities of the cascade's particles compensate for the different geometrical paths followed by the Cherenkov photons. The end result is that a Cherenkov shower is detected as a very narrow pulse of light.

will use an expression for the value of the refractive index, n, given in [320]:

$$n = 1 + 0.000292\left(\frac{x}{1030}\right)\left(\frac{273.2}{T}\right) \tag{4.7}$$

where $T = 204 + 0.091x$ K is the temperature as a function of the altitude, expressed with the atmospheric depth, x. Using Eq. 4.3 with $E_0 = 1$ TeV and $\langle E\rangle = 80$ MeV, we get an estimate of the atmospheric depth at the shower maximum for 1-TeV showers: i.e. $x = 506.2$ g cm^{-2}. So, the refractive index now equals $n = 1.00016$, and together with the derived value for β Eq. 4.5 yields $\theta \approx 0.9°$. Finally, using simple geometry the radius of the Cherenkov light pool at sea-level is

$$R = (8.4\ \text{km})\tan\theta \approx 138\ \text{m} \tag{4.8}$$

This value is slightly larger than that from Monte Carlo simulations ($R \approx 120$ m [1]). In reality, however, the energy dissipation during the development of an EAS follows an *exponential* law, rather than the one assumed in the above simple model, which leads to narrower emission.

Of course, the rest of the Cherenkov light–producing regions will also exhibit the same geometry, hence illuminating the area surrounded by the annulus with photons generated at altitudes below the shower maximum. Some illumination of the areas beyond the extent of the annulus will also occur due to Cherenkov production at altitudes above the shower maximum (Fig. 4.5). The integrated contribution, from all altitudes, to the Cherenkov light pool in the air is that of a cone which casts a Cherenkov shadow on the ground. The shadow's radius at sea-level extends to several hundred meters from the primary's projected impact point. However, the most luminous area is contained within R m.

In addition to its spatial characteristics, an EAS also exhibits unique temporal characteristics. The temporal spread of a Cherenkov light pool is very short (i.e. $\Delta t \sim 1$ ns). The latter statement can be quantified using first principles. However, such detailed analysis is outside the purposes of this thesis. Instead, one can use the simplified model of Fig. 4.6 to estimate the order of magnitude of the Cherenkov pulse duration:

> A Cherenkov wavefront produced (say) at time t=0 s, by charged $e^{-}-e^{+}$ pairs at the top of the shower maximum area (position O) will immediately begin

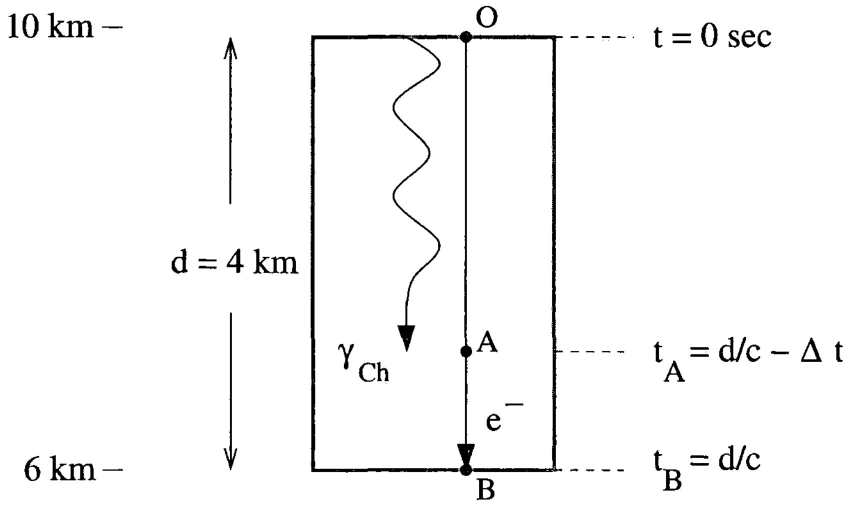


Figure 4.6: A toy model for the rough estimation of the Cherenkov flash duration. This simple geometrical representation assumes that the Cherenkov photons follow vertical paths to the ground. Cherenkov light emitted at the top (O) and bottom (B) of the shower maximum is strongly bunched in time because the particles which generate it travel close to the speed of light, whereas the Cherenkov photons do so at a slightly slower speed, $v_{\rm Ch} = c/n$. Hence, the time it takes for the particles — which travel at the speed of light — to cross the shower maximum (OB) is slightly shorter than that for the Cherenkov photons. At the moment when the particles arrive at point (B), the Cherenkov photons are behind by distance (AB).

its journey to the ground, at the speed of light in the *air* (i.e. $v = c/n$). The production of a second wavefront, this time from the lowest part of the shower maximum (position B), will have to wait until the charged particles reach that altitude. This can only happen after $\approx d/c$ s, since those particles travel close to the speed of light as they cross the longitudinal difference, d, between the lowest and the highest boundary of the shower maximum. At this point, the second wavefront starts making its way towards the ground. The time difference $\Delta t = t_{\mathrm{OB}} - t_{\mathrm{OA}}$ is the width of a Cherenkov light pulse and is given by the approximate formula

$$\Delta t \sim \frac{d}{v} - \frac{d}{c} = \frac{d}{c}(n-1) \approx 4 \text{ ns} \tag{4.9}$$

where n is assumed to be constant.

In this simplified model we have not considered the variation of n with altitude, the angle of emission, θ, and other parameters. In reality, the Cherenkov wave-fronts disperse through the atmosphere as they travel towards the ground, which causes the pulse width to broaden. The result is Cherenkov pulses $\approx$ 5–7-ns wide.

4.1.2 Hadron-Initiated EAS

In General

γ rays are not the only primaries capable of triggering EAS. In fact, they constitute the weak minority compared to the EAS caused by cosmic rays, which are naturally a lot more abundant than γ rays from even the strongest high-energy source we have ever observed. The cosmic-ray EAS can be divided into two types: those initiated by cosmic electrons and those, by the rest of the charged particle species: mainly protons. The first type is a lot less abundant than the proton-initiated EAS, but it can exceed the diffuse γ-ray background by as much as 100 times, depending on the measured value of the latter. On the other hand, proton-initiated EAS are 10^3 times more abundant than even the strongest γ-ray source detected [1]!

Nevertheless, a more useful categorisation exists if we consider the morphological and physical differences between the three types of EAS. Atmospheric cascades initiated by cosmic electrons are practically indistinguishable from those initiated by γ rays, and both are regarded as electromagnetic because their secondary products are either photons or electrons. The cosmic-electron cascade development is described by the same particle interactions, illustrated in Fig. 4.2 (see previous section), as for a γ-ray cascade.

On the contrary, proton-initiated EAS are different in nature, as they involve particle interactions that result in heavier secondaries. After a cosmic-ray proton has entered the Earth's atmosphere, it is very likely to collide with an atomic nucleus and thus break down to a variety of elementary particles and anti-particles. The general form of such an interaction is

$$p + N \longrightarrow n(\pi^{\pm}, \pi^{0}, p\overline{p}, \dots) \tag{4.10}$$

where N is the atmospheric nucleus.

In section 2.5, it was mentioned that the most common products of a cosmic-ray EAS are charged and neutral pions. If the secondary charged pions have $E > 100$ GeV, then it is likely that they will collide further to give even more of the above particle species. For lower energies, they simply decay to muons and anti-neutrinos. More concisely,

$$\pi^{\pm} + N \xrightarrow{E>100\mathrm{GeV}} n(\pi^{\pm}, \pi^{0}, p\overline{p}, \dots) \tag{4.11}$$

$$\pi^{\pm} \xrightarrow{E<100\mathrm{GeV}} \mu^{\pm} + \overline{\nu}_{\mu} \tag{4.12}$$

Furthermore, the neutral pions produced from the hadronic interaction have very short lifetimes and decay to two γ rays, which can generate their own EM cascades. These cascades develop as part of the total hadronic EAS and cannot be distinguished from the latter.

Finally, the produced muons from the above interactions have also short half-lives — although much longer than those of the pions — and decay to e^{-}/e^{+} and neutrinos according to

$$\mu^{\pm} \xrightarrow{\tau_{1/2}\approx 1.4\ \mu\mathrm{s}} e^{\pm} + \overline{\nu}_{\mu}/\nu_{\mu} + \overline{\nu}_{e}/\nu_{e} \tag{4.13}$$

Their production from pions occurs only at low altitudes, below the shower maximum ($\ll 6$ km), which characterises them as "local". Local muons can trigger air showers whose light pools have limited spread, due to the low altitude at which they develop. In addition, their non-negligible half-lives allow them to traverse, at nearly the speed of light, a distance $c\tau_{1/2} \approx 400$ m before they decay. This property makes them potentially detectable from the ground using counter arrays or water-pond detectors.

A schematic of the full-cascade interactions in a proton EAS can be found in Fig. 4.7.

Main Properties and Differences from EM EAS

There are a few distinct differences between EM and hadron-initiated showers. To begin with, it is clear that hadronic EAS involve particle interactions during the cascade, which are different from those in EM EAS. Those interactions involve inelastic collisions between hadrons and atmospheric nuclei, which result in the fragmentation of the latter into lighter nuclei with significantly larger transverse momenta than the EM products from Bremsstrahlung and pair production. Consequently, the lateral extent of a hadron-initiated shower is broader. In the case of EM showers, the average opening angle between the produced electrons and the shower's direction is [321]

$$\langle\theta\rangle = \frac{m_e c^2}{E_e} \text{ rad} \tag{4.14}$$

where $m_e c^2$ is the electron rest energy, and E_e is their relativistic electron energy. For 1-GeV primaries this leads to $\theta \approx 0.2^\circ$ [1]. Hence, a γ-ray-initiated EAS is tightly bunched in the forward direction. In addition, hadronic showers can be a lot more penetrating in the atmosphere and have a longer tail than EM ones, although on a shower-to-shower basis they exhibit larger fluctuations with regards to longitudinal development (see Fig. 4.8).

Hadronic showers exhibit a wide variety of secondary products, like $\pi^\pm$, $\mu^\pm$, $e^\pm$ and $p\bar{p}$, only a small part of which are electromagnetic during the shower's first stages of development. Hence, if one compares hadronic and EM showers of equal primary energies, the former are less efficient at producing Cherenkov light. The reason for this is that particles like $e^\pm$ and $\mu^\pm$ which are efficient at emitting Cherenkov light are only produced in substantial quantities at later stages of the shower's development, at which point, of course, their energies are well reduced with respect to those of their progenitors. Typically, the light output from γ-ray showers is 2–3 times higher than that from protons of the same energy [1]. Moreover, as the energy of the primaries decreases, hadronic EAS become much more inefficient at producing Cherenkov light than showers produced by γ rays of the same energy.

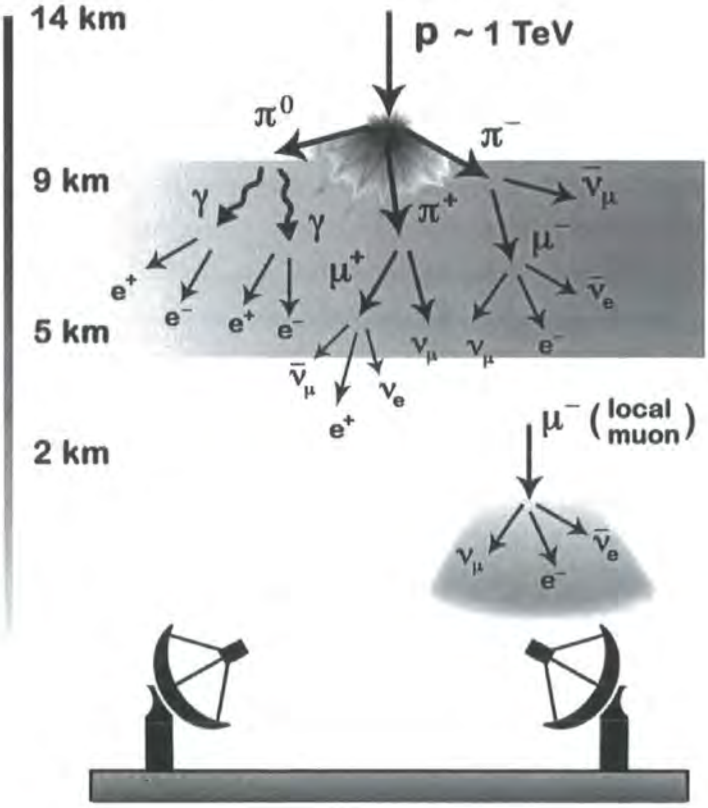


Figure 4.7: The interactions of 1-TeV proton-initiated EAS. In such cascades the emission angles of the different interactions are much broader than those in EM cascades. Furthermore, the proton showers develop further down in the atmosphere than their EM equivalents and give rise to a local muon component which triggers low-energy, localised air showers. The shower maximum for 1-TeV proton showers (shaded region) ranges from 5 to 9 km a.s.l. (compare with Fig. 4.2). This type of EAS occurs 10^3–10^4 times more frequently than the γ-ray-initiated cascades.

The bunching in time that is apparent for EM showers is not as well defined as for hadronic ones. Instead, their temporal width is somewhat longer. The main reason for this is the local muon component caused by superluminal muons that reach the ground much earlier than the Cherenkov photons from the shower maximum. As a result, the Cherenkov emission from hadronic showers displays precursor pulses, which prolongs the overall emission.

A basic property of hadronic EAS, which was unsuccessfully used as a discriminant in first generation γ-ray experiments, is their *isotropic* nature. The charged cosmic ray particles are largely affected by the interstellar magnetic fields, which conceals any information on the former's origin. Hence, the arrival directions of cosmic-ray showers are isotropically distributed across the sky. On the contrary, γ-ray showers from high-energy sources are expected to point back to their origin and thus show a fair amount of anisotropy. Although it sounds plausible to utilise such a technique for γ/hadron discrimination, unfortunately the technology of those early experiments required very large γ-ray fluxes — a few percent of the cosmic-ray background — for a source to stand out in the hadronic background.

In this paragraph we discussed the basic differences between γ-ray and hadronic EAS. However, the crucial matter of discriminating between the two in an effort to detect and pinpoint γ-ray sources from the ground is the subject of the following paragraphs. In those paragraps, the Imaging Atmospheric Cherenkov Technique and its benefits from stereoscopic VHE observations are discussed.

4.2 Collecting All That Light

It seems appropriate to use Cherenkov light from γ-ray EAS to do astronomy. The secondary products of a γ-ray cascade do not deviate a lot from the shower axis, and the Cherenkov emission angle is quite narrow, so that the showers can be traced to their source. Yet the Cherenkov light pool on the ground extends several hundreds of metres away from the shower axis, which means that even the smallest of detectors are likely to detect part of a shower's light. In addition, the amount of light that every shower produces is directly proportional to the number of particles in the cascade [323] and can be, therefore, used to estimate the energy of the primary. Finally, the very short duration of a Cherenkov signal makes it unique amongst other sources of light, like starlight and meteor showers, and can be used as a discriminant against them.

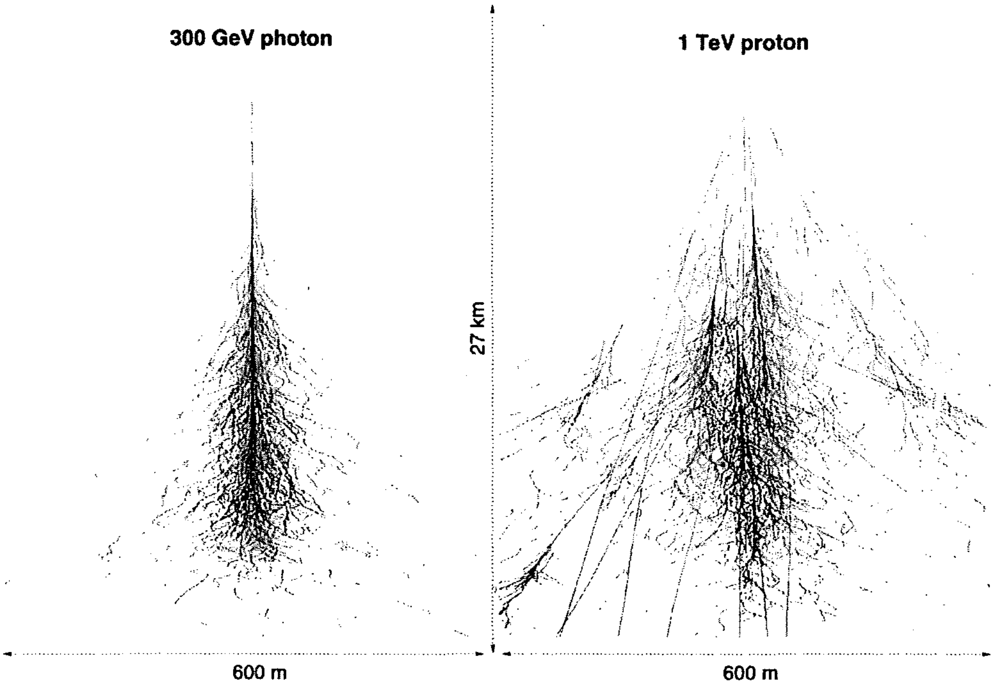


Figure 4.8: (from [322]) Monte Carlo simulations of the longitudinal and lateral development of 1-TeV proton and 300-GeV γ-ray showers. The traces represent the paths of the shower particles as they travel through the atmosphere.

The detection of a shower's Cherenkov light from the ground does not necessarily require elaborate detector systems. The number of EAS generated every second in the Earth's atmosphere is enormous. For example, modern experiments like H.E.S.S. can detect roughly 1,000 Cherenkov events every second, above 100 GeV. In principle, atmospheric Cherenkov emission can be detected with a simple system that involves, for example, a 2-m^2 reflector with a fairly high reflectivity of $R \approx 85\%$, and a data acquisition system with short integration times of the order of 10 ns [1]. Such elementary systems were used in the pioneering experiments by Galbraith and Jelley in 1953 and provided the first evidence for ground-based observability of high-energy emission (see section 1.3.3) [30].

4.2.1 Predictable Factors

Hadronic Background

The key problem in doing ground-based γ-ray astronomy lies in the method of recognising the few γ-ray showers against a large background of hadronic ones. The early experiments did not have the ability to make this distinction. Initially, most of the efforts to detect γ-ray sources relied upon the anisotropy of the amount of showers across the emission area, as opposed to the isotropic background signal elsewhere. The Lebedev Institute made the first serious attempt to try and utilise this method towards the detection of the Crab nebula. Unfortunately it was without success, as their sensitivity was below the required (see section 1.3.3).

Still, there are other shower properties one could exploit in order to achieve a discrimination. It was mentioned amongst the properties of a hadronic EAS that there is a local muon component which is responsible for the wider spread of Cherenkov light. These muons reach the ground and can be potentially detected. Unfortunately, it often happens that the muon component is missing, which makes the use of such information as a discriminant unreliable. Alternative factors, like the polarisation of the Cherenkov light from EAS, have also been considered, but, again, it has not been proven practical. Instead, the discriminant that led to the detection of numerous γ-ray sources and allowed us to do efficient γ-ray astronomy came from the consistency of the geometrical shapes and directions of the γ-ray showers. The method is described in section 4.2.4.

Muon Background

Local atmospheric muons are the products of hadronic EAS and their Cherenkov emission can trigger a telescope. Since their energies are much lower than those of their primaries, the muon background becomes important only for detectors with low energy thresholds. As it will be discussed later, they can be easily rejected with stereoscopic observations (see section 4.2.6).

Night Sky Background

Another obstacle which could potentially hinder ground-based γ-ray observations is the fluctuations of the Night Sky Background (NSB) light. Starlight fluctuates much more slowly than the typical duration of a Cherenkov flash ($\Delta t_{\mathrm{Ch}} \approx$ 5–7 ns). Its intensity, however, sets a limit to how faint the Cherenkov events should be in order to still be detected. The primary energy threshold, above which the telescope can operate without being unacceptably triggered by starlight, depends on the distribution of the night-sky fluctuations and that of the Cherenkov events — which are mostly of hadronic origin. Its value within the wavelength sensitivity bounds of a detector's PMTs, λ_1 and λ_2, is proportional to

$$E_{\mathrm{th}} \propto \frac{\sqrt{B(\lambda)}}{C(\lambda)} \sqrt{\frac{\omega \tau_{\mathrm{w}}}{\eta(\lambda) A}} \tag{4.15}$$

where $C(\lambda)$ and $B(\lambda)$ are the Cherenkov and NSB photon fluxes within the sensitivity bounds; $\eta(\lambda)$ is the response curve of the PMT and ω the solid angle of the detector; A is the mirror collection area of the detector, and, finally, τ_{w} is the time window over which the camera's PMTs integrate light, and it is typically longer than a Cherenkov flash [1].

Fortunately, the NSB levels can be measured and removed from the observed data with a process called *image cleaning.* The process involves recording the NSB pedestal (i.e. the NSB levels) in parallel with an observation, for each pixel on the camera, and then subtracting it from the data. In general, however, the integration time of the camera ($\sim$ 10 ns) is optimised towards minimising the accidental triggers due to the slow NSB fluctuations.

Finally, if there are bright stars in the FoV during an observation, their position is taken into account, and the random events triggering the PMTs in that direction are discarded. A detailed description of the estimation of the NSB for the H.E.S.S. experiment can be found in [324].

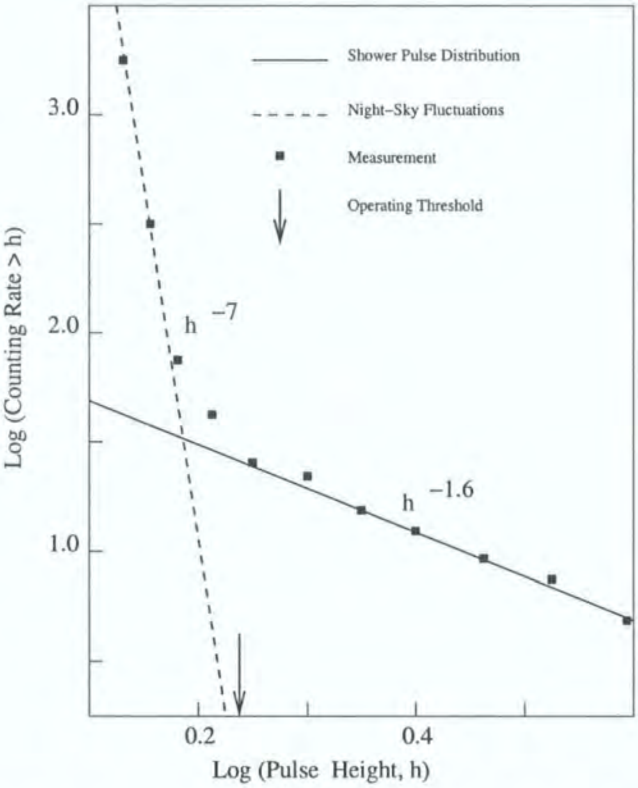


Figure 4.9: (from [1]) Pulse-height distribution of telescope events during an observation. The low-energy part of the distribution is dominated by the NSB spectrum, which is soft, whereas above a certain pulse height — or equivalent event energy — the distribution becomes hard and clearly dominated by Cherenkov events. Note that the energy threshold is defined by the position on the spectrum where the triggers from NSB events become sparse.

Electronics

Even when an Air Cherenkov telescope is not exposed to the natural light sources, its PMTs are affected by electronic noise, called *dark pedestal*. The event distribution during darkness follows a Gaussian whose shape and position depends on the temperature as well as the voltage gain applied to the PMTs. The locations where most ground-based experiments are based at have temperatures which typically vary by $\delta T = 20°\mathrm{C}$, over a yearly observation period. Therefore, the dark pedestal has to be constantly checked.

The Moon and Other Light Sources

Although γ-ray astronomers select the experiment locations very carefully, unwanted light emission may occur because of various reasons, which include transiting aeroplanes, meteor showers, city lights, airport beacons, etc. Although seemingly impossible to get rid of, the accidental triggers from these sources can, again, be minimised using short camera-integration times.

However, if any of those sources becomes brighter than the operational limits of the PMTs, then they could potentially cause damage to the sensitive electronics. A clear example of such a source is the moon. Observation planning depends largely on prediction tables of moonrise and moonset, so that permanent damage to the PMTs can be avoided (see Fig. 4.10). Unfortunately, having to observe only during completely dark nights reduces the observation time by significant amounts.

4.2.2 Unpredictable Factors

The Atmosphere

Despite the general success in minimising the factors that can hinder γ-ray observations, there are still others that cannot be controlled. The most unpredictable of all is also the reason for the existence of this branch of astronomy: the atmosphere. The Earth's atmosphere provides us with a vast, replenishable amount of active gas, with which particles and photons interact. (The secondary products from these interactions are directly or indirectly detectable and reveal the presence of high-energy particles in the atmosphere.) It is a natural calorimeter that has protected us from deadly cosmic and γ rays, and which has only recently become part of our detectors. Any other astronomer would have been grateful, or at least unconcerned, if their

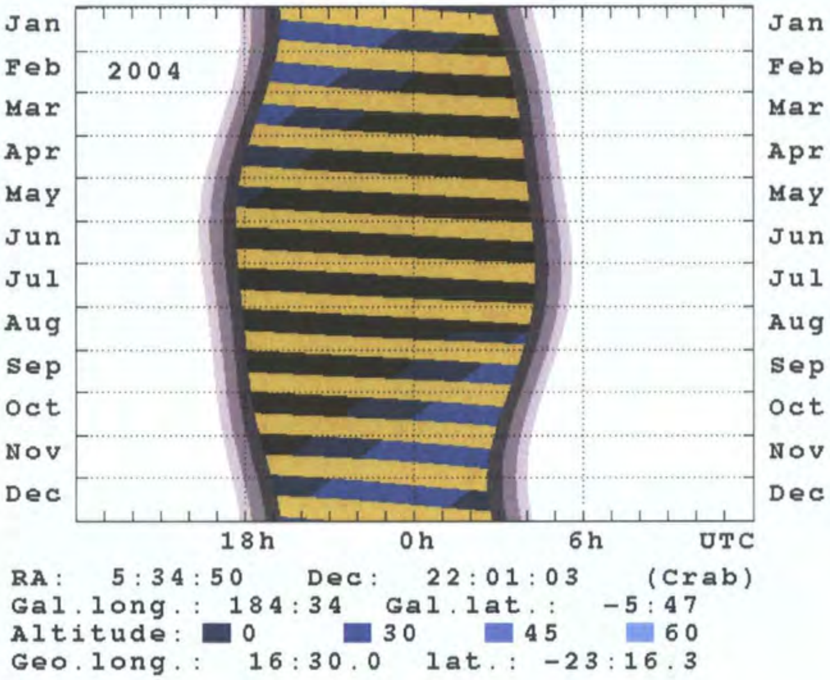


Figure 4.10: (from [325]) Visibility of the Crab nebula from the H.E.S.S. site in Namibia, for 2004. The dates on the horizontal axis are given in UTC time starting at midnight. The white area around the coloured zones corresponds to sunlight, and the grey zones to the various twilight definitions. The yellow bands across the coloured zones correspond to moonshine. The graduated blue colours correspond to the time periods when the Crab is above the given altitudes, as seen from Namibia. Finally, in the black area the object is below the horizon and, therefore, invisible. Such plots assist the observation schedules for specific sources and allow astronomers to programme their upcoming yearly observations.

observatories (and not themselves!) were placed in space. For a VHE astronomer the atmosphere is an essential element.

The atmosphere is also an unavoidable obstacle because it is a dynamical system which constantly changes. Temperature, humidity, atmospheric extinction, etc. are all parameters that depend on time. The way Cherenkov light propagates through the different layers of the atmosphere is largely affected by these parameters and makes the measured shower energies on the ground a varying quantity, even for identical showers. At the moment, it seems very difficult to correct for these second-order effects. However, most modern experiments are now equipped with weather instruments that can measure all the critical quantities, like humidity and temperature, as well as atmospheric transmissivity and scattering characteristics. These measurements are then used in the post-processing of the data, in combination with standard, theoretical atmospheric profiles. This helps to assess the energy and direction of the primaries more accurately. The instruments that are most useful for such measurements are the weather station, the ceilometer and the infrared radiometer. These instruments are now widely used in collaborations like H.E.S.S., Fly's eye, Pierre Auger, etc. A brief description for each instrument is given in section 4.3.

Cosmic Electron/Positron Background

Triggers from EM showers caused by cosmic electrons (or positrons) are impossible to reject because their shower development is identical to that of γ rays. Luckily, their energy spectrum steepens beyond 100 GeV (see Fig. 4.11), which results in an insignificant contribution to the overall *trigger rate* (the strict definition of this term is explained later in this section). As expected, the electron/positron component does not show preferred direction, so a large part of it can be rejected by excluding off-source events.

4.2.3 Trigger Rate and Effective Area

For every Cherenkov detector, the detectable number of γ rays per unit time (dN_γ/dt), from a source, divided by the total γ-ray flux of that source at Earth, F_γ, defines the detector's *effective area*, A_{eff}. In other words, A_{eff} is the area across which the detector can effectively detect those events. The lateral spread of the Cherenkov light pool at sea-level means that the mirror area of a Cherenkov detector does not need to be large for the experiment to achieve an efficient detection

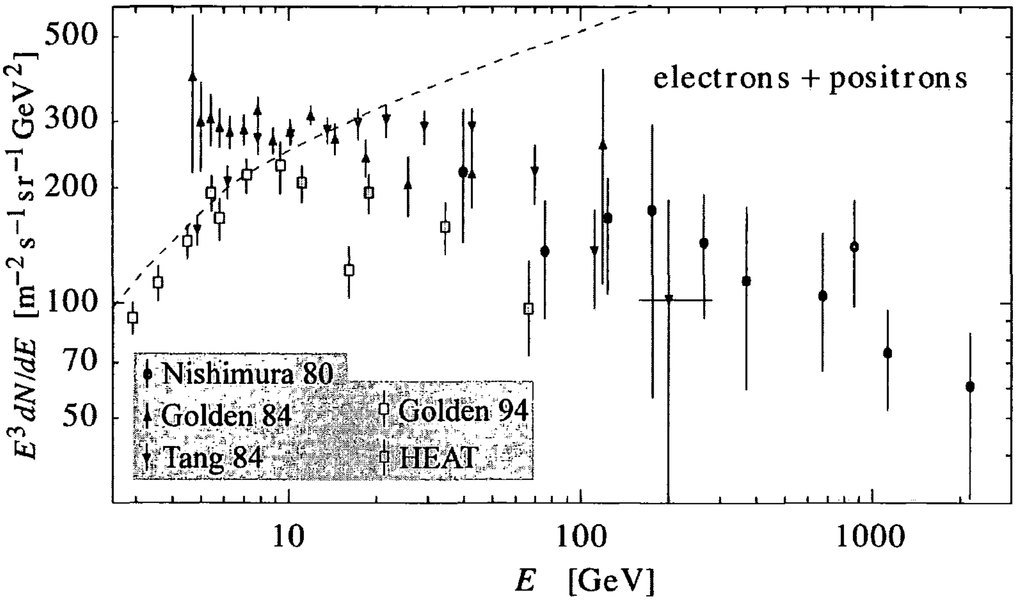


Figure 4.11: (from [326]) Measured cosmic electron and positron spectra (E^3dN/dE) with various balloon experiments [327],[328],[329],[330],[326]. The dashed line represents the cosmic proton spectrum multiplied by 0.01.

rate, otherwise known as trigger rate. Nevertheless, the larger the reflector, the higher the probability that the amount of collected light will be detected above the background noise. This probability depends on the distance R between the detector and the impact point of the shower axis on the ground, the relative orientation between the incident-shower direction and the telescope axis, θ, and the shower energy, E. In addition, because in general the response of a detector varies within its FoV, and because the Cherenkov light that reaches the detector is not uniformly distributed across the FoV, there is a dependence on the solid angle across which the shower maximum is viewed from the telescope [321]. A graphical representation of the above parameters is shown in Fig. 4.12.

The general form of the detection probability function is

$$pr(E,\theta,r,\omega) = \frac{N_{\text{trig}}(E,\theta,r,\omega)}{N_{\text{tot}}(E,\theta,r,\omega)} \tag{4.16}$$

where $N_{\text{trig}}(E,\theta,r,\omega)$ is the number of triggered showers with energy E and direction θ that land at distance r from a detector with FoV ω; and $N_{\text{tot}}(E,\theta,r,\omega)$ is the total number of such showers. Across the whole FoV of a detector the effective area is then

$$A_{\text{eff}}(E,\theta) = 2\pi \int_{\omega} \int_{0}^{\infty} pr(E,\theta,r,\omega) r dr d\omega \tag{4.17}$$

In practice, its value is calculated with Monte Carlo simulations by simulating a large number of showers with different energies, directions and distances from the detector, and counting the number of triggers. It can be seen that beyond a specific radius, r_0, the number of triggers becomes insignificant. Fig. 4.13 shows the case where the incident angles were fixed to vertical incidence and only 10 GeV–10 TeV showers were considered. In that case, nearly all triggers were confined within 400 m distance from the detector. By performing such simulations, one can set the maximum area, πr_0^2, which will contain the total number of simulated showers. Having fixed r_0, the effective area can be calculated from Eq. 4.17 as a function of the shower energy, E, and the incident shower angle, θ. For showers with energies and directions in the intervals $[E, E+\Delta E]$ and $[\theta, \theta+\Delta\theta]$, respectively, this equation becomes

$$A_{\text{eff}}(\Delta E,\Delta\theta) = \omega\pi r_0^2 \cdot \frac{N_{\text{trig}}(\Delta E,\Delta\theta)}{N_{\text{tot}}(\Delta E,\Delta\theta)} \tag{4.18}$$

where we assumed that the detection probability remains constant across the FoV. The size of the energy and direction intervals can be chosen arbitrarily in simulations. It makes sense, however, to make them comparable to the telescope's energy and angular resolutions.

The energy that corresponds to the point where the effective-area function starts declining, as the detector loses sensitivity, is defined by the detector's *energy threshold* (see Fig. 4.14). For experiments with high energy resolutions (e.g. $\Delta E/E = 20\%$), these plots are especially useful because they can be used to predict the expected number of events above a certain energy threshold, with small statistical errors. Nevertheless, different definitions of the energy threshold can be used depending on the particularities of the analysis (see section 5.7).

Provided there is a known value of the effective area for different energies, one can calculate the *differential* trigger rate (i.e. particles per time and energy interval) for a detectable source of primary particles. If the particles are distributed according to a single differential spectrum $dN(E)/dE$, then the differential rate is

$$\frac{dR(E)}{dE} = \frac{dN(E)}{dE} \cdot A_{\text{eff}}(E) \text{ particles s}^{-1} \text{ GeV}^{-1} \tag{4.19}$$

If the spectrum of the particles is described by a single power law over an energy range $[E_0, E_2]$, then one can integrate Eq. 4.19 to get the *integral* trigger rate, $R(E_0 \rightarrow E_2)$:

$$R(E_0 \rightarrow E_2) = \int_{E_0}^{E_2} \frac{dN(E)}{dE} \cdot A_{\text{eff}}(E) dE \text{ particles s}^{-1} \tag{4.20}$$

A more precise definition of a detector's energy threshold is the value of E which corresponds to the maximum differential rate. The existence of such a maximum is justified by the opposite gradients of the effective-area function below the experiment's threshold and the typical γ-ray spectra in the VHE region (see section 3.2.2 and Fig. 3.26). Provided there is an energy overlap between the two competing functions, the resulting differential rate has the shape shown in Fig. 4.15.

Upper Limits on the Flux

The expected trigger rate for a particular source over an energy range $[E_0, E_2]$, as defined by Equation 4.20, can be used to set upper limits on the the source's differential and integral fluxes, in the case of non-detection. By means of statistical tests (e.g. the χ^2-test), some of which are described in section 5.4, one can give an

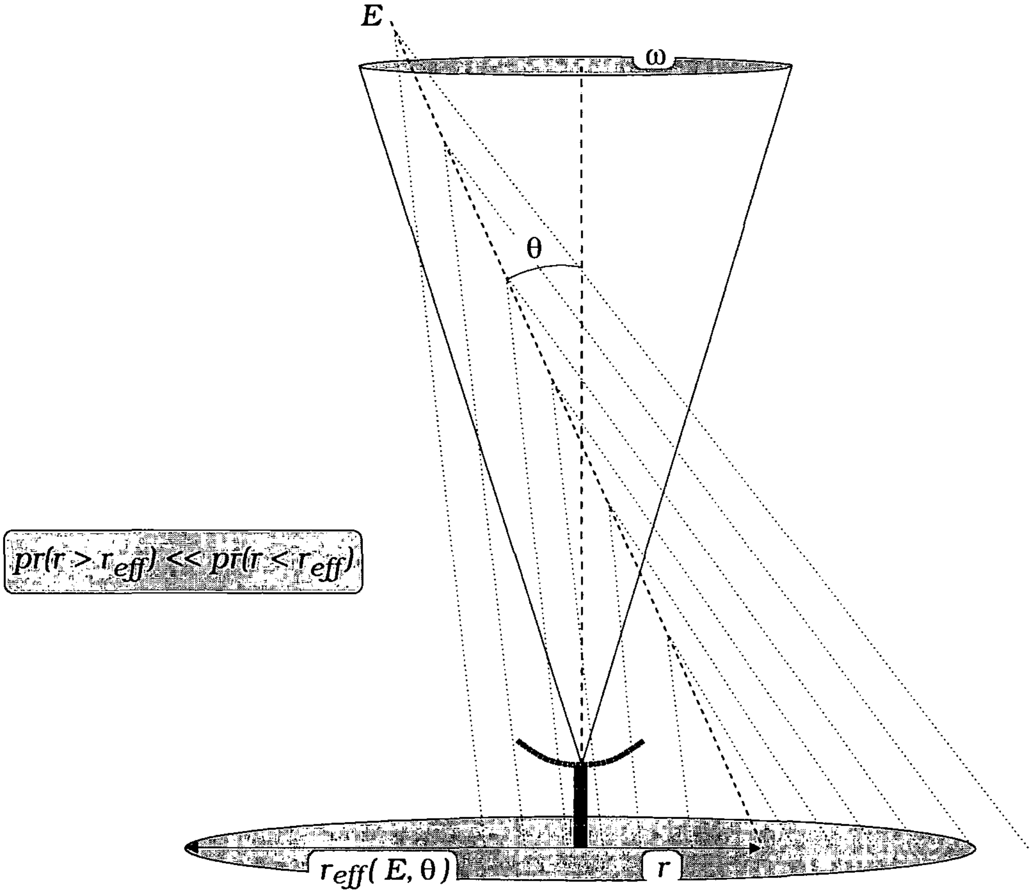


Figure 4.12: The definition of the parameters on which the probability of detecting an EAS, $pr(E, \theta, r, \omega)$, are dependent. In this figure, an EAS of energy E lands at distance r from a telescope. Apart from those two parameters, the probability of detecting such a shower also depends on the telescope's FoV, which is defined by the solid angle, ω, and the relative orientation between the telescope's axis and the shower's direction, θ. Beyond a certain distance, r_{eff}, the probability of detecting showers of specific θ decreases dramatically. This limiting distance defines the effective area of the telescope. In the present figure, we have assumed azimuthal symmetry for the detector's response across its FoV. In general this need not be true, and the generalised form of the effective area is given by Eq. 4.17.

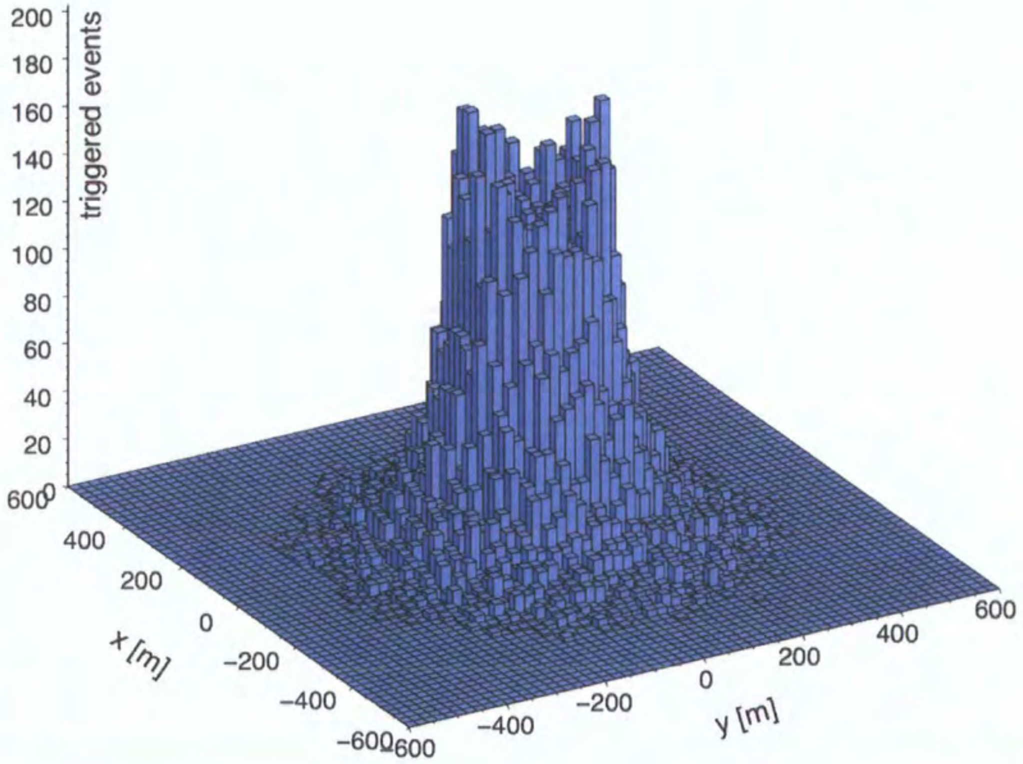


Figure 4.13: (from [321]) Monte Carlo simulation of the dependence of the number of triggers on the shower distance from the detector, for a single telescope. The total number of simulated showers is 10^6, the shower energies are in the 10 GeV–10 TeV range, and the shower directions are vertical to the ground. It can be seen that beyond $r_0 = 400$ m the number of triggers becomes insignificant.

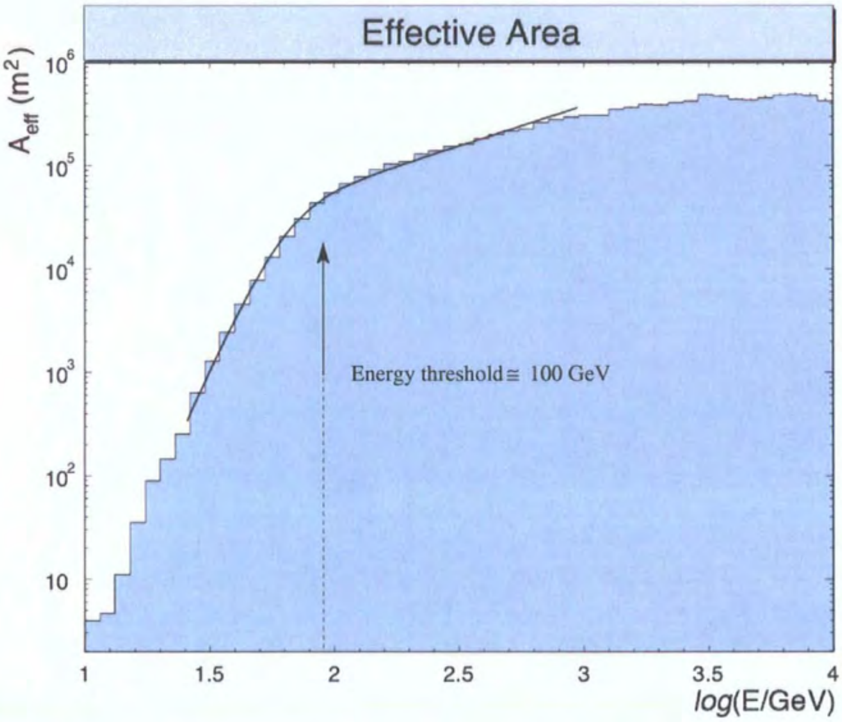


Figure 4.14: (from [321]) The variation of A_{eff} with energy, for 10 GeV–100 TeV, vertical showers. The energy range has been divided into intervals of $\Delta \log(E/\text{GeV}) \approx 0.7$, and the effective area has been calculated for each interval. The steep decline of the effective area below $\approx$ 100 GeV shows the threshold behaviour of the detector.

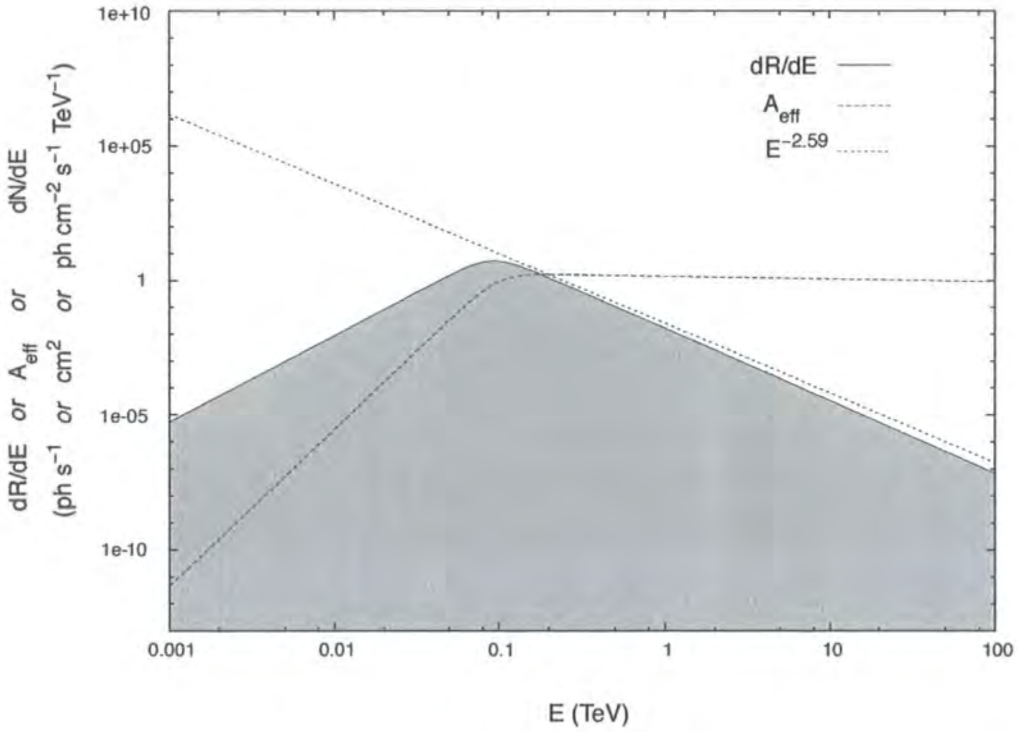


Figure 4.15: The typical shape of a differential-rate plot (solid line), where the maximum defines the energy threshold (E_{th}) of the experiment for a specific source with a differential spectrum $dN/dE \propto E^{-\nu}$. In this particular case the spectral index was chosen equal to -2.59. The effective-area (dashed line) and power-law functions (doted line) have been rescaled in order to match as closely as possible the shape of the differential-rate function at energies below and above the energy threshold, respectively. In this plot, $E_{\mathrm{th}} \approx 90$ GeV.

estimate on the number of signal events that would need to have been present above the background noise, in order to be considered as a clear signal at a given level of significance (e.g. 3 Poissonian standard deviations). If for an observation that yielded N_{ON} on-source and N_{OFF} off-source events this number is $N_{3\sigma}(N_{\mathrm{ON}}, N_{\mathrm{OFF}})$ — when a 3-σ confidence level (C.L.) is required — and if the exposure time of the observation (sometimes referred to as *livetime*) is T_{live}, then the upper limits on the differential and integral flux are

- Differential flux upper limit at $E = E_1 \in [E_0, E_2]$

$$\left[\frac{dN(E_1)}{dE}\right]_{\mathrm{ul}}^{3\sigma} = \frac{N_{3\sigma}(N_{\mathrm{ON}}, N_{\mathrm{OFF}})}{T_{\mathrm{live}} \cdot R(E_0 \to E_2)} \cdot \frac{dN(E_1)}{dE} \ \mathrm{cm}^{-2}\ \mathrm{s}^{-1}\ \mathrm{GeV}^{-1} \tag{4.21}$$

- Integral flux upper limit across $[E_0, E_2]$

$$F_{\mathrm{ul}}^{3\sigma}(E_0 \to E_2) = \frac{N_{3\sigma}(N_{\mathrm{ON}}, N_{\mathrm{OFF}})}{T_{\mathrm{live}} \cdot R(E_0 \to E_2)} \cdot \int_{E_0}^{E_2} \frac{dN(E)}{dE} dE \ \mathrm{cm}^{-2}\ \mathrm{s}^{-1} \tag{4.22}$$

Note that the differential flux, $dN(E_1)/dE$, is always calculated across an infinitesimal energy range, $(E_1, E_1 + dE)$, which is centred, in this case, on an arbitrary value, E_1, that belongs to $[E_0, E_2]$. On the other hand, the integral flux, $F(E_0 \to E_2)$, is by definition calculated across the whole interval, $[E_0, E_2]$, which makes it a lot more meaningful, since the energies of the primaries cannot be restricted to infinitesimal ranges. Typically, $[E_0, E_2]$ is chosen so that it covers the whole energy spectrum in which a detector can operate reliably. Hence, a lower limit corresponding to the detector's energy threshold for the particular data set is usually set. A limit at higher energies is not usually needed, because the fluxes diminish long before the energy limit up to which most Cherenkov detectors are optimised ($\approx$ 50 TeV [1]). More discussion about flux upper limits takes place in chapter 6, where we apply the above formulae to the data from three pulsars: the Crab, PSR B1706−44 and PSR B1259−63.

Clearly, the calculation of upper limits using the above expressions requires knowledge of the source spectrum. Since the latter is not generally known in the energy range of interest — otherwise one would not be restricted to upper limit calculation — it has to be assumed. These assumptions are typically based on observations in different energy bands, but also on theoretical modelling. Thus, it can be said that the derived upper limits are *model-dependent.*

4.2.4 The Imaging Atmospheric Cherenkov Technique

A Cherenkov telescope collects the light from EAS with a large reflector and redirects it onto a 2-dimensional array of PMTs, i.e. the telescope's camera. In other words, it projects the 3-dimensional structure of a shower onto a 2-dimensional surface. The shape and intensity of this projection, which is called the *shower image*, depends on a few critical factors:

Geometry

Cherenkov telescopes are not wide field, survey-type instruments. In the vast majority of cases, knowledge of the source position is required prior to a detection — although serendipitous sources have also been detected with Cherenkov telescopes (see e.g. [134] and [133]). Therefore, if during an on-source observation the telescope's optical axis is aligned with the direction from which a γ-ray EAS is expected, then the following are true:

- All paraxial showers have images whose shapes are roughly elliptical and point to the camera centre. In the rare case of a perfect match between the shower and telescope axes, these images are circular and located at the camera centre.
- Showers that arrive at random angles with respect to the telescope's pointing direction have images with no preferred orientation.

The image of a shower is the superposition of light generated at various stages of a cascade's development. For a shower whose axis is parallel to the telescope's, but whose *core location* (i.e. its impact point on the ground) is as much as $\approx$ 120 m away from the telescope — assumed to be at sea-level — both the lateral and the longitudinal extent is projected on the camera (see Fig. 4.16a,b). Since the former is substantially broader than the latter, the projected shapes are the recognisable elliptical shapes of Cherenkov showers. Furthermore, the part of the image that appears closer to the centre of the camera is due to Cherenkov light which was generated at the top of the cascade. Light from the shower maximum forms the bulky, central region, whereas the least energetic particles at later stages of the cascade's development contribute to the far end of the image. Finally, the more distant the shower's core location, the more elongated the images.

Physics

The most important factor that influences the image shape is the type of cascade that generated the incident light. As was mentioned in previous sections, hadronic showers generally have a broad and rather erratic development. Their images on the telescope's camera are themselves broad and, since hadrons are isotropically distributed, randomly orientated on the camera. On the contrary, γ-ray images are more consistent with elliptical shapes than hadronic ones, which also makes their orientation better defined. For γ rays that originate from a source in the telescope's FoV, all images should point towards the source position on the camera plane (e.g. the camera centre).

Additionally, a Cherenkov telescope can also be triggered by local muons that generate light low in the atmosphere. The corresponding images appear as rings (or part of a ring) on the camera and are usually referred to as *muon rings*. The three types of Cherenkov images are shown in Fig. 4.17.

It becomes clear that the angular distribution of the image orientation can be a very powerful discriminant. But how is this quantified in order to optimise the percentage of γ-ray showers in the data? In the next section, we present an important formalism that allows us to not only discriminate between cosmic- and γ-ray events, but also to deduce some basic properties of the shower from its recorded image.

4.2.5 The Image Parameters

In 1985, A.M. Hillas presented a parametrisation scheme for the collected shower images, which helped to assign specific acceptance ranges to the image orientation and shape and, thus, lead to better discrimination between hadronic and γ-ray showers [332]. According to this formalism, the Cherenkov images recorded on the camera can be described by an ellipse that best represents the image shape which is formed by the registered light on the PMTs. The geometrical properties of these best ellipses are widely known as *Hillas parameters*, named after the person who first proposed their usability, and can be directly connected with a shower's orientation, lateral and temporal extent, etc. Fig. 4.16c shows the geometrical definition for each of these parameters.

As mentioned earlier, the most important discriminant between hadrons and γ rays is the image orientation with respect to the telescope axis. This quantity is

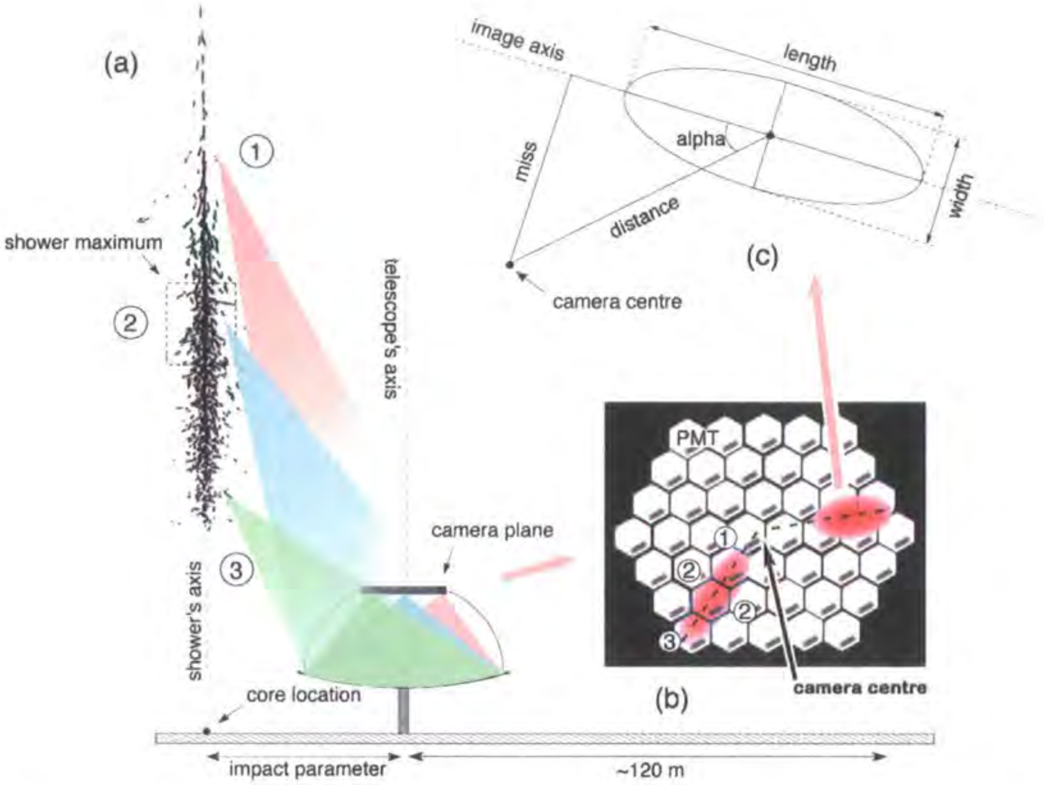


Figure 4.16: Recording the geometrical features of an EAS. (a) As the shower develops in the atmosphere, Cherenkov light from the different parts of the cascade is emitted towards the telescope. (b) If the shower lands as much as 120 m away from the telescope, which is assumed at sea-level, both the lateral and the longitudinal extent of the shower is recorded in the form of elliptical images on the telescope's camera. The longest dimension of these images corresponds to the longitudinal spread, (1)–(3), whereas the dimension perpendicular to that corresponds to the lateral spread, (2)–(2). Light from the top of the shower (1) is projected closer to the camera centre than that from the bottom (3). The central part of the shower image is formed by Cherenkov emission from the shower maximum (2). (c) After the shower image has been collected by the telescope, it is fitted with an ellipse whose properties, called the image or Hillas parameters, carry information about the shape and position of the shower relative to the detector. These parameters help us discriminate between hadronic and γ-ray showers. The various elements of this figure, i.e. telescope, shower, etc., are not to scale.

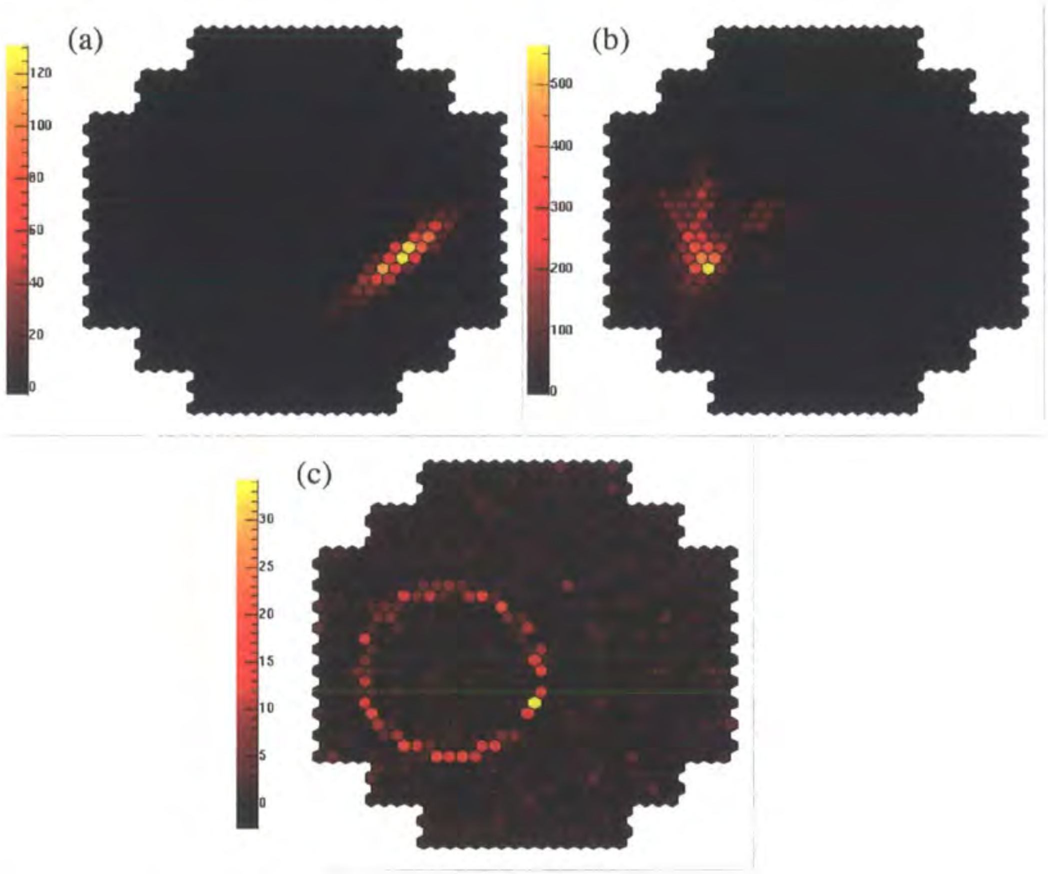


Figure 4.17: (from [331]) Three types of Cherenkov images: (a) a γ-ray event, (b) a hadron event and (c) a muon ring. The colour variation corresponds to light intensity.

expressed with the *alpha* parameter, which is equal to the angle — measured in degrees — between the semi-major axis of the ellipse and the position vector that connects the centre of the camera with that of the image. For sources that are centred on the telescope's FoV, the *alpha* parameters of the ellipses are close to zero and increase with increasing shower inclination with respect to the telescope axis.

The length of the position vector of the ellipse corresponds to the parameter *distance*. Since the area of the camera plane corresponds to the telescope's FoV, which covers a certain angular distance across the sky, this parameter expresses the angular distance (usually measured in radians) of an image from the centre of the camera. In terms of its connection with the shower properties, the *distance* parameter carries some information regarding the radial distance of the core location from the telescope, also known as the *impact parameter*. However, it does not provide an 1–1 relation with it. Strictly speaking, its value is equal to the angle between the position vector of the shower maximum, as seen from the telescope, and the shower axis. Therefore, the *distance* parameter is a very useful measure of how far from the telescope the shower landed, but it does not give the exact distance to the core location, which is something that can be determined with better accuracy in stereo observations, as we shall see later (see Fig. 4.18).

The lateral and longitudinal spread of a shower is proportional to the *width* and *length* of the ellipse, respectively. In geometrical terms, the *width* and *length* are equal to the semi-minor and semi-major axes, respectively. The further a paraxial shower lands from the telescope, the more elongated its image becomes, which means that *length* is increased.

Finally, the parameter *miss* expresses the angular distance between the camera centre and the expected source position. If the latter coincides with the camera centre, then both *alpha* and *miss* should be close to zero for events coming from the source direction and should have larger values for off-source events. As Fig. 4.16 shows, the *miss* parameter is an alternative measure of the image orientation, and its value is minimised for showers arriving from the source position.

The Image Amplitude and Low-Energy Observations

The amount of light that is collected from each shower can be expressed in photoelectrons (ph.e.). On average, one ph.e. is generated when a photon ionises the gas of one of the telescope's PMTs. The ph.e. are then accelerated across a high electric potential (~ 1 kV) and generate electrical signals, whose strength is represented

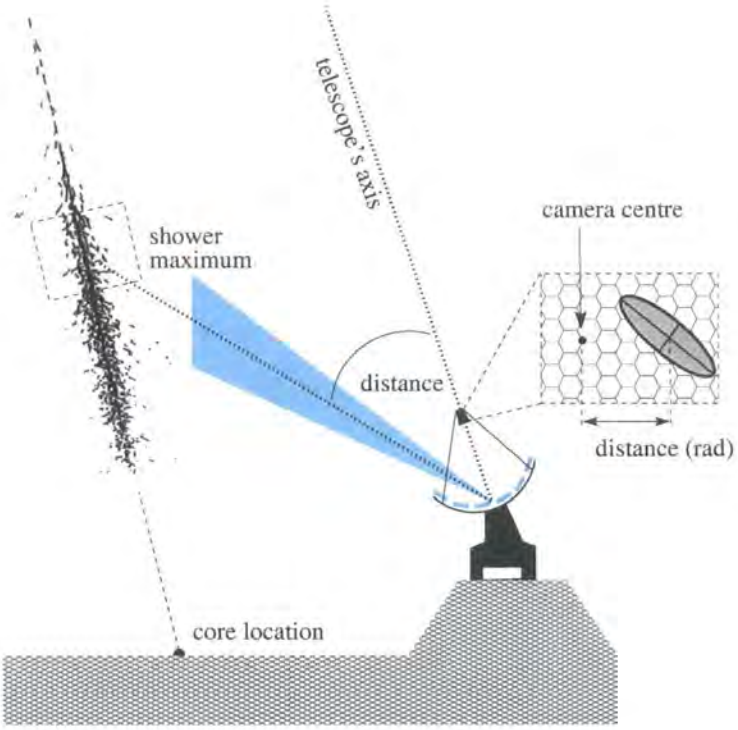


Figure 4.18: The geometrical definition of the *distance* parameter. The light blue-shaded section represents the uncertainty in determining the distance to the shower maximum.

by a number of digital counts (d.c.). More about the function of PMTs and other components of a Cherenkov telescope will be discussed in section 4.3.

Every image recorded contains a certain amount of ph.e. Fig. 4.17 shows such an image, in which there is a number of pixels that are brighter than others. This means that those pixels have generated more ph.e. during their integration time. In γ-ray images, there is a distinct intensity gradient as one moves from the image centre towards its edge, which forms a narrow ellipse. If we add all the ph.e. from all the pixels on the camera after having subtracted the electronic pedestal together with the NSB noise, we obtain the image *size* of an event, also known as its *Image Amplitude* (*IA*). Division of the number of ph.e. in the two brightest pixels of the image by the *IA* gives the *concentration* of a Cherenkov event. The first parameter is an indication of a shower's energy, E, whereas the second corresponds to the compactness of a shower and shows how broad the energy distribution along the various stages of a cascade is.

Unfortunately, the *Image Amplitude* alone cannot supply the energy of the primary. By simulating numerous γ-ray showers of various energies with a Monte Carlo code, one can construct a relation between *IA* and E for a wide range of γ-ray energies. It turns out that shower images from showers with different energies can give the same amount of ph.e. on the camera, and showers of the same energy can appear to have different *Image Amplitudes*. The reason for these differences is the uncertainty on the core location, which has to be restricted if one is to successfully correlate *IA* with energy. Without knowledge of the distance between the light source and the telescope, there remains a possibility for close-by, low-energy showers to produce the same amount of detectable light as more distant, high-energy ones. The uncertainty involved in the *IA*–E correlation is strengthened with decreasing *IA*, as Fig. 4.19a shows, and the reason is the still detectable — but severely attenuated — light from a substantial amount of distant, high-energy showers which appear as 'small', in terms of *size*, on the camera. On the other hand, if one samples only showers whose *distances* are below a certain value, e.g. 18 mrad, then the linear correlation becomes more evident and the high-energy outliers are practically eliminated (see Fig. 4.19b).

The constraints that we place on the image parameters in order to create a sample of shower images sharing similar features, like *IA*, orientation, etc., are called *image cuts*. The use of image cuts is usually towards enhancing the γ-ray signal over the hadronic/muon background: for example, the application of a combination of cuts

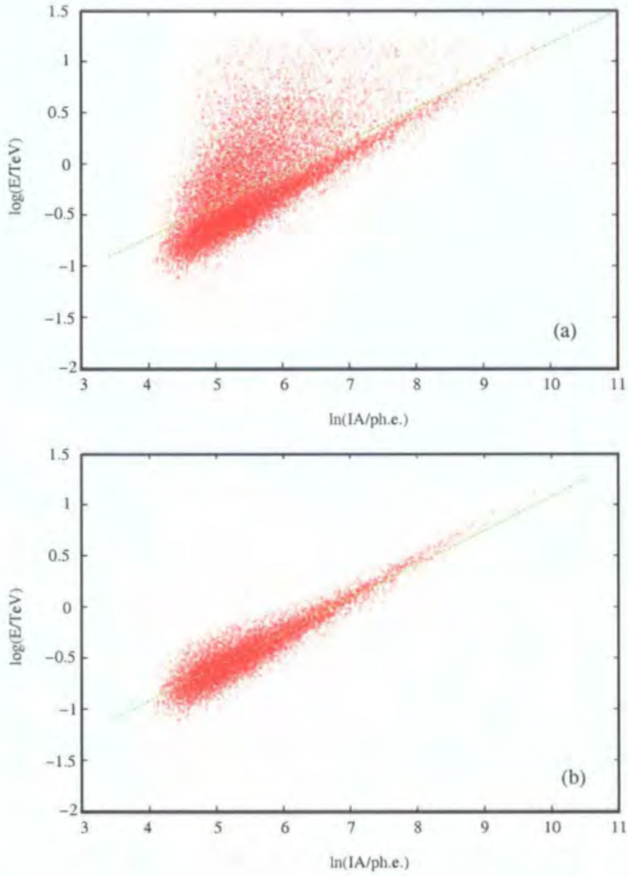


Figure 4.19: (a) Correlation plot between *Image Amplitude* (*IA*) and event energy, E, from a Monte Carlo simulation of 20,000 γ-ray showers. Although there is a clear linear dependence between the parameters for most of the showers, the fact that the core locations of the simulated showers were randomly scattered around the telescope resulted in a noticeable spread towards the higher energies. As the number of ph.e. decreases, the amount of showers departing from the linear law increases. The green line is a fit on the whole data and is clearly biased by the low-*IA* spread. (b) The same correlation for events with *distances* ≤ 18 mrad: i.e. more central on the camera plane. Linearity is now better resolved. The bulge at the lower energies is due to the uncertainties which are involved in the determination of a shower's impact parameter for low *IA*.

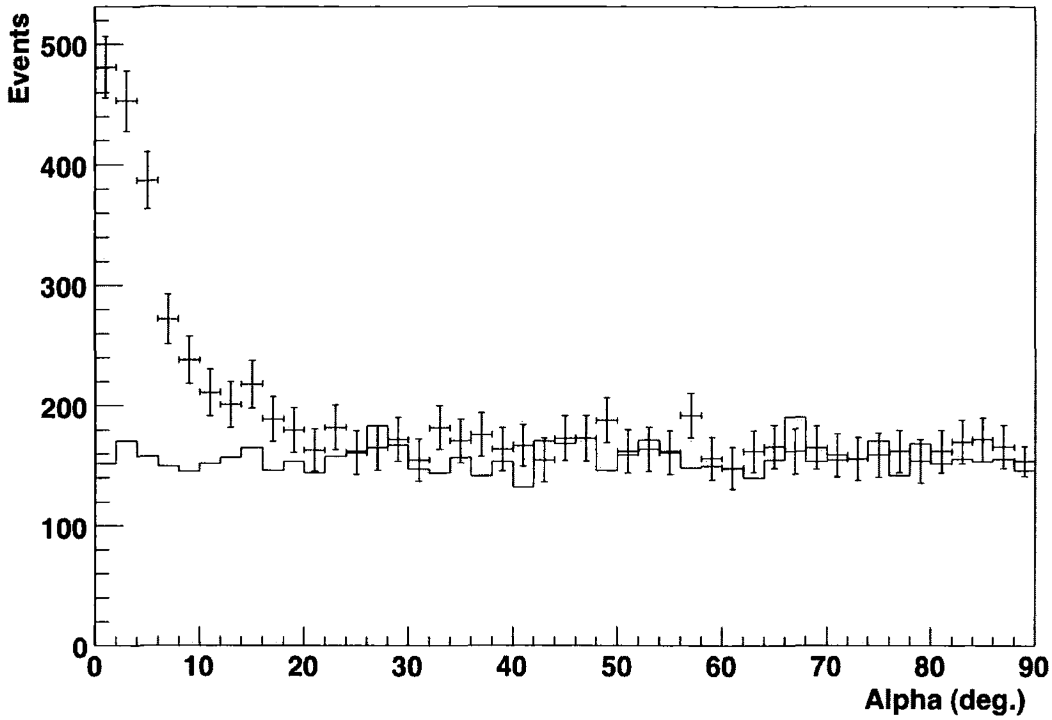


Figure 4.20: (from [334]) *Alpha*-plot for on- and off-source, single-telescope observations of the Crab nebula. The randomly orientated hadronic images contribute towards a flat off-source distribution, whereas the directional consistency of γ-ray images results in a steep rise of the on-source distribution, for low *alpha* angles.

to the shower image parameters can eliminate as much as 99.98% of the background, while still retaining a respectable amount of γ rays ($\approx$ 30%) [333]. Fig. 4.20 shows an example of the effectiveness of an additional *alpha* cut on the events that have passed the rest of the cuts. It can be seen that the off-source event distribution (histogram), which contains mainly hadronic events, is flat, whereas γ-ray events that are included in the on-source observation (crosses) contribute towards a steep rise below a certain *alpha* angle. By placing an upper limit on the *alpha* parameter somewhere around 9°, one can maximise the significance of the γ-ray signal.

Another useful image cut, the *length/size* (*LoverS*), can help reduce the local-muon background events that typically plagues low-energy, GeV observations. In general, the population of muon showers has low surface brightness, which is characteristic of their low energies. As a result, the muon rings on a telescope's camera have small *IA*. In addition, their characteristic ring shapes make them less compact than

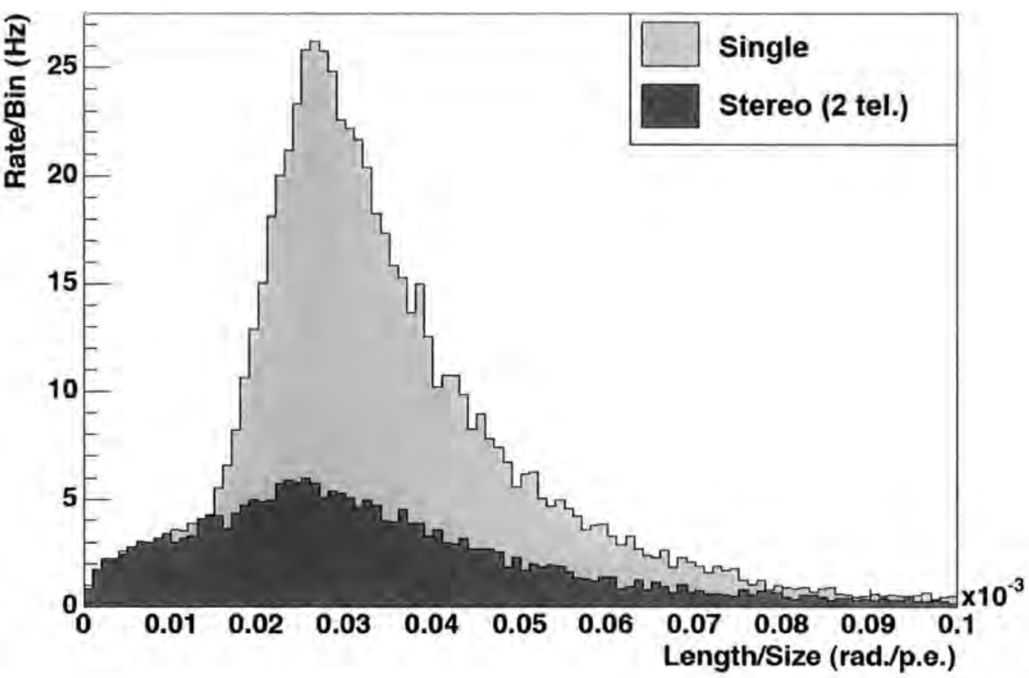


Figure 4.21: (from [335]) Distribution of the *LoverS* parameter for single (light grey) and 2-telescope, stereo (dark grey) observations. The narrow peak of the distribution in single-telescope observations is due to the local muon component. A constraint on the values of *LoverS* below ~ 0.018 mrad ph.e.$^{-1}$ rejects virtually all muon events. On the other hand, stereo observations achieve muon rejection by exploiting the local character of these events, which do not trigger the telescopes coincidentally.

the average γ-ray- and proton-induced showers, which results in larger associated image lengths (see Fig. 4.17). As a consequence, muon images have large values of *LoverS*, which can be confined to small values through the application of a cut to the data, in order to reduce the muon background (Fig. 4.21). Finally, in stereoscopic observations the muon background is reduced by means of a coincidence-trigger requirement between two or more telescopes that are spatially separated; so, a cut on *LoverS* is not necessary (see section 4.2.6).

When applying cuts to the data, one has to make sure they are meaningful: most of the image cuts become unreliable below a certain *IA* because there are not enough pixels in the shower's image for its parameters to be clearly defined. The process of assigning image parameters to Cherenkov events is called *image reconstruction*. Image cuts associated with orientation and shape are likely to fail for these images

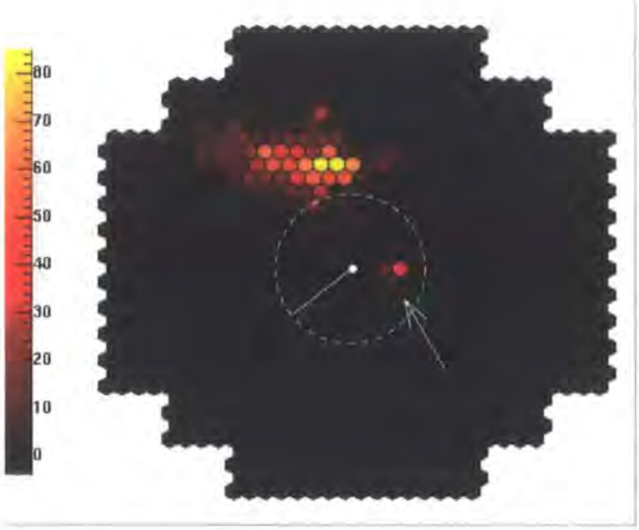


Figure 4.22: (from [331]) The appearance of small events close to the centre of the camera, an example of which is shown with the help of the white arrow, is connected with low-energy showers. One can select such events by setting an upper limit on the *distance* of the recorded images. In this image, we have superimposed a circle (dashed line) whose radius corresponds to the upper limit on the *distance* parameter, in order to highlight the area that contains the selected events. This way, it is possible to perform a low-energy analysis targeting at pulsar GeV emission.

and result in wrong judgement of the shower's origin. However, one can still restrict the *distance* for such small events in order to keep only those which appear close to the camera centre (see Fig 4.22). This way it is possible to exclude small events originating from energetic but distant showers and, hence, set an upper limit on the energy of those showers. Concentrating only on low-energy events can be very useful for the detection of pulsed, GeV emission from pulsars. Unfortunately, the lack of full image reconstruction means that the background is not rejected as effectively, which decreases the significance of a possible detection. Moreover, off-source events have all possible directions and energies and can easily generate small, centralised shower images which contaminate the desired sample. Fortunately, for low-energy, pulsar observations, one can exploit the periodic nature of the emission, which can be used as a discriminant against the random background.

4.2.6 Principles of Stereoscopic Observations

The simultaneous observation of Cherenkov events with more than one telescopes, otherwise known as *stereoscopy*, was pioneered by the H.E.G.R.A. collaboration (see section 1.3.4) and has been widely adopted by follow-up experiments, like H.E.S.S. and VERITAS. There are significant benefits to be gained from stereoscopic observations, which mainly arise from the 3-dimensional information that can be inferred about the shower by observing it from different locations.

The simplest stereoscopic systems have two detectors, say CT1 and CT2, separated by an optimal distance which allows for the best coverage of the Cherenkov light pool at the altitude of the experiment. For experiments built close to the ground, a separation of $\sim$ 100 m is typically used. This value can vary depending on the altitude of the experiment, so that the Cherenkov light pool is efficiently covered [35]. Each telescope records the image of a shower as if it were stand-alone. However, the events that trigger the whole system as one (typically coded CT0) have to satisfy certain criteria. Fast electronics implemented for such systems are able to decide on-the-fly which showers will be recorded based on these requirements. The most common requirements that the events have to satisfy are the following:

- **Temporal coincidence.** The events recorded by CT1 and CT2 have to register with CT0 within a specific time window. If only one of the telescopes records an image within the specified time window, then the event is discarded. For configurations with telescopes 100 m apart, this window is typically $\sim$ 2 μs.

  Prior to deciding for the distances between the individual telescopes of a stereoscopic array, the temporal coincidence criterion is taken into account so that the system trigger rate is optimal. Systems that are well-spread apart are less likely to be triggered by a single shower than those that are close-packed. However, the former systems will have better stereoscopic view, if triggered, which allows for better localisation of the showers, as we shall see later.

- **Topological coincidence.** The images recorded by each telescope should be spatially correlated for CT0 to register a shower event. In other words, the displacement of an image relative to another on the camera should be proportional to the telescope separation.

These need not be the only criteria for stereoscopic triggering, and depending on the situation even just the first condition could provide reliable EAS detection. The reader can find an in-depth justification of the above, as well as additional trigger strategies in [336].

After the events from stereoscopic observations have been recorded, the analysis process combines the images from the two telescopes towards a better shower localisation; fig. 4.23 shows how this is done. The impact parameter for a shower can be determined by the intersection of the semi-major axes of the shower images that each telescope has recorded. Furthermore, the orientation of showers that arrive at an angle relative to the telescopes' axes can be calculated by superimposing the images from both cameras (planes R1 and R2) on a reference plane (R3), i.e. R1+R2→R3, and intersecting their semi-major axes. The angular distance between the intersection points and the centre of the camera, d, corresponds to the inclination of the shower's axis relative to that of R3. This distance can be translated into inclination via the focal length of the telescopes: i.e.

$$\theta = \frac{d}{f} \text{ rad} \tag{4.23}$$

where f is the focal length of the telescope.

The θ parameter in stereo observations is the equivalent to the *alpha* parameter in single-telescope observations, since it reveals the direction of the shower relative to the telescopes' axes. However, contrary to *alpha*, it is a lot less ambiguous since any directional uncertainty of the images is practically removed via triangulation. Having in our hands such a powerful parameter, it becomes obvious that the selection requirements should aim at constraining this parameter to small values.

Although the principle of stereoscopy can be achieved with at least two detectors, the core location of faint showers can be hard or impossible to determine because their image axes are not well defined. Hence, the overdetermination of the shower location with three or more telescopes is common practice because it increases the positional accuracy even further and allows the localisation of even fainter showers [35]. In these cases, the intersection of the three or more images rarely leads to a single point on the reference plane, but rather to a confined area. Therefore, it is much more useful to define a parameter which will be able to divide the camera's FoV into small and large inclination sections. This parameter is the θ^2 and it defines a circular section around the expected shower position: e.g. the centre of the camera. Showers whose images intersect inside an area defined by the maximum cut on θ^2

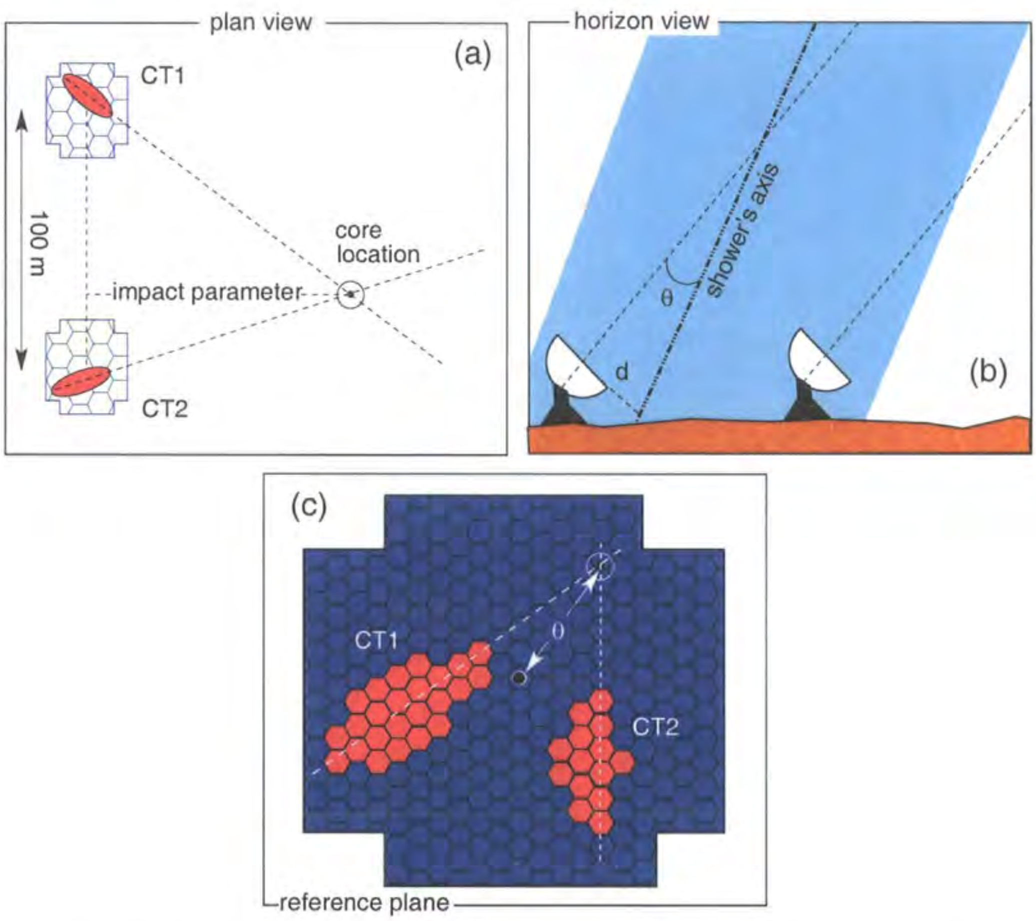


Figure 4.23: The principle of determining the shower location and direction with a simple stereo system. (a) The impact point of the shower on the plane perpendicular to telescopes' axes can be determined by intersecting the semi-major axes of the recorded images from each telescope. (b) In the case where the incident shower forms an angle θ with the telescopes' axes, we can determine its value by superimposing the images from both showers onto a reference camera plane (c). The angular distance from the intersection point of the images' semi-major axes to the centre of the camera is proportional to the angle θ.

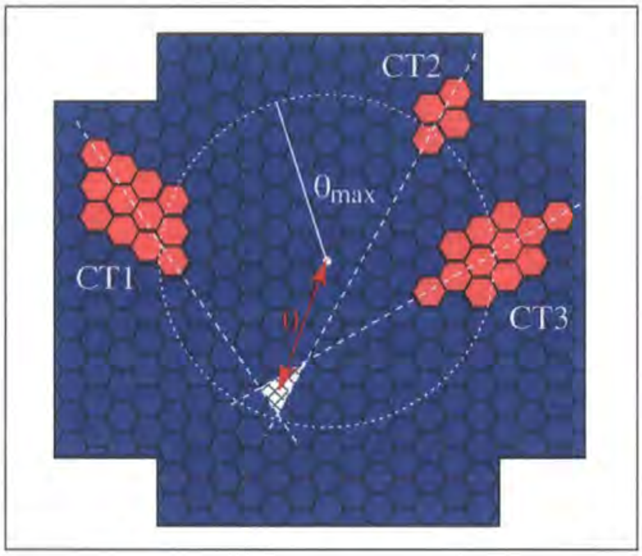


Figure 4.24: Superposition of the images of the same shower, as recorded by three Cherenkov telescopes: CT1, CT2 and CT3. The θ^2 cut would select such showers, only if their images intersect below a certain distance from the expected shower direction — in this case, the centre of the camera.

are accepted as coming from γ rays, provided, of course, they have passed the rest of the cuts. Fig. 4.24 shows an example of such scenario.

Stereoscopic systems have significantly lower thresholds than their stand-alone counterparts. This arises from a combination of factors. To begin with, the use of more than one telescopes means that the overall collection area is increased, which allows for a greater percentage of a shower's Cherenkov light to be collected, and therefore makes it possible to detect fainter showers. However, the contribution of the increased collection area towards a lower energy threshold is not as significant as if it were used with a single, large detector. More importantly, the energy threshold reduction with stereoscopic systems arises from:

- **Stereoscopic reconstruction.** The ability to determine the position of the showers with higher accuracy results in better discrimination between γ rays and hadrons and can still be effective at low energies, where single-telescope images are too dim to resolve the orientation of the showers.
- **Multi-telescope trigger.** Large Cherenkov detectors have low energy thresholds, and their low-energy triggers are mainly due to local atmospheric muons

and the NSB. When operating in coincidence, these triggers can be very effectively rejected using the fact that muon showers and NSB fluctuations are unlikely to trigger multiple telescopes. This allows the detection of EAS in the energy region where typically a single telescope would be too dominated by NSB fluctuations to operate effectively.

Moreover, the energy and angular resolution is enhanced with stereoscopic systems:

- **Improved energy resolution.** The use of two or more telescopes for the estimation of a shower's impact parameter reduces dramatically the errors involved in the calculation of the shower energy. By using a suitable averaging algorithm, one can fit the different intersection points and resolve the shower direction and core location [337]. Modern experiments, like H.E.S.S., can achieve an energy resolution of $\Delta E/E \approx 20\%$ per event, for a wide range of energies [35]. The ability to estimate the shower energy more accurately is very useful for spectroscopy and results in better correlation between the image *size* and energy.
- **Improved angular resolution.** Determination of the shower direction with a resolution better than 0.1° is now possible with stereoscopic systems [35]. This is clearly an asset when observing point sources like pulsars, as it can help to reject γ-rays from nearby sources whose images point away from the expected position. It is also a valuable tool in the efforts to resolve extended sources: e.g. the PWN MSH 15−52 and the SNR SN 1006 [67],[83].

Recent successes with stereoscopic observations include the resolution of the supernova shell RX J1713.7−3946 at TeV energies, and the detection of the binary pulsar PSR B1259−63, along with the unidentified, serendipitous source HESS J1303−631 (see Fig. 4.25): HESS J1303−631 appeared in the same FoV, $\approx 0.6^\circ$ north of the pulsar, and analysis of data taken in February and June 2004 showed a significant excess over the background of 21 σ [73],[117],[133]. More discussion on PSR B1259−63 and its surroundings takes place in section 6.3.3.

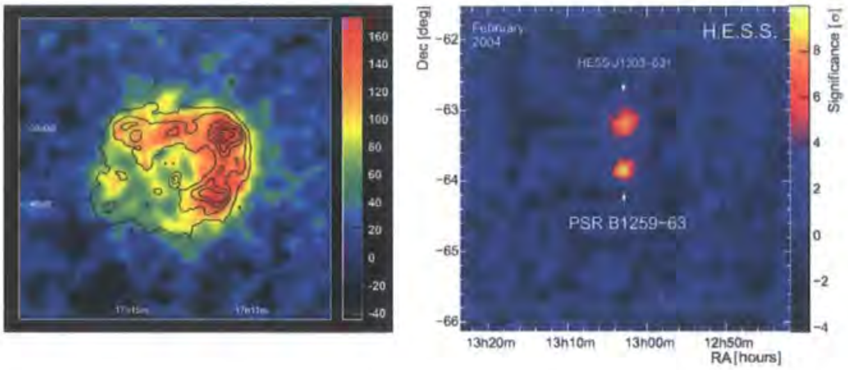


Figure 4.25: (from [73] and [117]) Sky maps of the TeV supernova remnant RX J1713.7−3946 (left) and the binary pulsar PSR B1259−63 with the serendipitous, unidentified γ-ray source HESS J1303−631. The contours on the left plot show the X-ray surface brightness as it was derived from *ASCA* observations in the 1–3 keV range. Significance is represented by the colour gradient: e.g. the significance of PSR B1259−63 in the plot on the right is 9.1 σ. Source position is expressed in RA–Dec co-ordinates (x–y axes). The maps have been smoothed such that each pixel's intensity is represented by a Gaussian with standard deviation equal to the instrument's resolution for this data set. For the left plot, this corresponds to 3 arcmin.

4.3 The High Energy Stereoscopic System: From Concept to Realisation

In March 1997, the Max Planck Institute for Nuclear Physics (MPIK) in Heidelberg revealed their intentions for a construction of a new stereoscopic array of Cherenkov detectors [338]. The motivation for such a project arose from the already established H.E.G.R.A. experiment, which had been successfully producing results up to then, but which was also near completion. The suggested name for this large-scale project, which involves many international scientific groups, was H.E.S.S. (High Energy Stereoscopic System) — named after Victor Hess, the discoverer of cosmic radiation. The project's goal was to exceed H.E.G.R.A.'s sensitivity by approximately one order of magnitude and, hence, become capable of detecting a minimum energy flux of $F(>1\ \mathrm{TeV}) \sim 10^{-13}$ erg s^{-1} after 100 h of exposure time. This would allow to probe deeper into the non-thermal universe and increase by a statistically significant amount the detected high-energy sources with previous experiments. Fig. 4.26 compares the sensitivity of H.E.S.S. with the fitted IC spectrum of the Crab nebula from EGRET and other ground-based observations. The standard-candle sensitivity of a single, H.E.G.R.A.-type Cherenkov detector, which is equivalent to $\approx 20\%$ of the Crab nebula's flux at ≈ 1 TeV, is also presented in this plot. Finally, for comparison with future space-borne experiments, the sensitivity range of NASA's *GLAST* has been included.

In June 2002, the first telescope (Fig. 4.28) of the High Energy Stereoscopic System (H.E.S.S.) was erected on the Khomas Highland of Namibia ($16^\circ\ 30'\ 00''$ W, $23^\circ\ 16'\ 18''$ S) (see Fig. 4.27). The location provides excellent sky clarity around the year, and its high altitude (1,800 m a.s.l.) makes it suitable for Cherenkov-light detection [339]. Furthermore, the choice of a southern observatory meant that objects like the Galactic centre, as well as most of the EGRET pulsars, could now be observed at conveniently small Z.A.s, without their signal being severely attenuated during its journey through the atmosphere. Observations with the single Cherenkov detector commenced soon after its construction was completed, and by the end of 2003 all four telescopes became operational, which finalised this way H.E.S.S. Phase I (see Fig. 4.29). As of 2005, H.E.S.S. members comprise groups from Armenia, Czech Republic, France, Germany, Great Britain, Ireland, Namibia and the Republic of South Africa.

In the next sections, the various components of the H.E.S.S. telescopes as well as their combined operation are introduced.

4.3.1 Technical Aspects of the System

Telescope Arrangement

The 4 telescopes of H.E.S.S. Phase I are arranged in a way which maximises the stereoscopic view, but still retains acceptable trigger rates (see section 4.2.6). Each telescope is built on the corners of a square pattern with 120-m side length and with the square's diagonal orientated along the North–South direction (see Fig. 4.30).

Telescope Mount and Drive

The H.E.S.S. telescopes operate on an Altitude–Azimuth (Alt–Az) mount, which supports their large and heavy, steel structure. The azimuthal and altitudinal motion is performed on horizontal and vertical circular rails, respectively. The rails are ≈ 14 m in diameter and their circular shape minimises the drive forces acting upon them during motion (see Fig. 4.31 and Fig. 4.32). Having an Alt–Az mount means

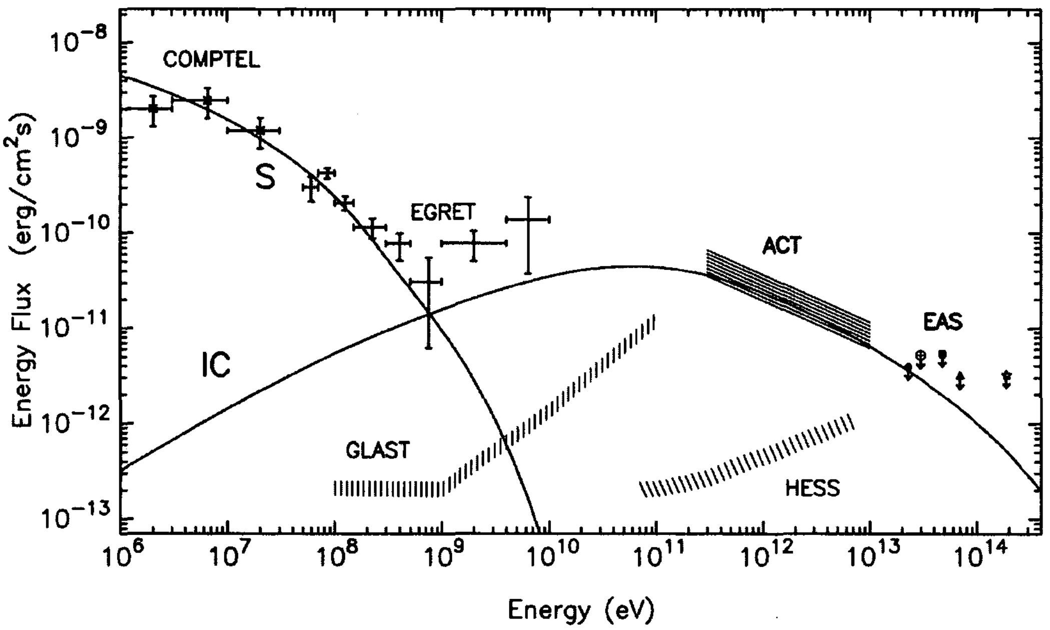


Figure 4.26: (from [338]) H.E.S.S.'s sensitivity across the energy range 80 GeV–10 TeV. For comparison, the plot shows model curves for the unpulsed inverse Compton spectrum of the Crab nebula (IC) and the pulsed synchrotron spectrum (S). The *CGRO* observations below 10 GeV are shown with plain error bars, and the TeV upper limits from ground-based experiments are shown with symbols on the far right of the plot. In addition, this plot has included the standard sensitivity of a single, Cherenkov telescope (ACT), i.e. ≈ 0.2 Crab, and that of NASA's future space-borne experiment, *GLAST*. Above 1 TeV, H.E.S.S. is expected to achieve sensitivities ~ 1 mCrab, i.e. 1/1,000 of the Crab nebula's flux in the same energy range.

Figure 4.27: (from [340]) The H.E.S.S. location in Namibia.

Figure 4.28: (from [331]) The first H.E.S.S. telescope. Each mirror element is 60 cm in diameter.

Figure 4.29: (from [331]) The full stereoscopic array of H.E.S.S. Phase I.

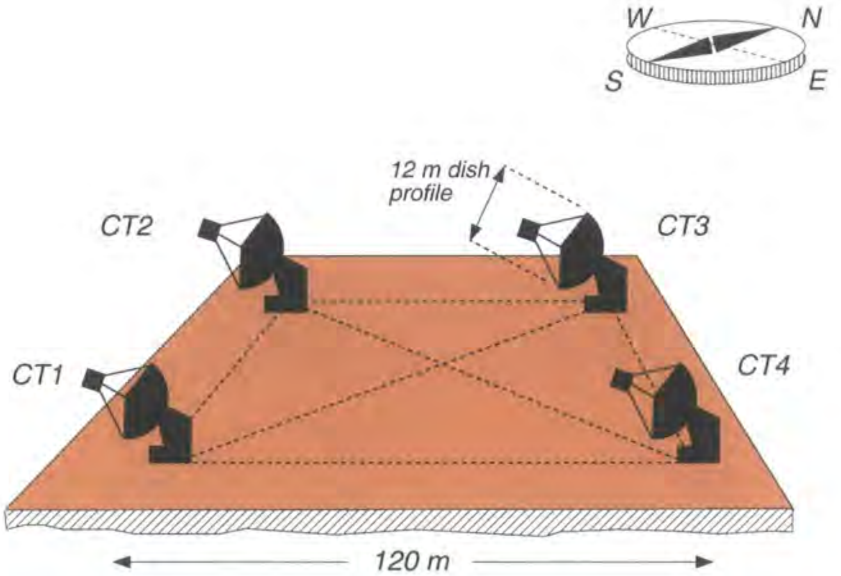


Figure 4.30: The telescope arrangement of H.E.S.S. Phase I. The 4, currently operating telescopes of Phase I (CT1–CT4) are arranged in a square formation with 120-m side length. Each telescope's dish is 12-m wide (flat-to-flat).

that star tracking cannot be controlled by just a single rotor. The complex calculations needed for the tracking across the celestial sphere are handled by a computer that controls the motion on both rails. In addition, the pointing is monitored and corrected with analogue and digital encoders connected to both rails. The resulting accuracy is of a few arcseconds, and it is further improved with the use of an optical guide telescope ($f = 800$ mm) that is equipped with a CCD camera (Sky CCD). Finally, the tracking speed can reach 100°/min at its maximum, which allows to readily point the 60-ton telescopes in the direction of a GRB, for example, in case H.E.S.S. has been notified for such occurrence. The layout of the H.E.S.S. telescopes is described in [341].

Mirror Design

The H.E.S.S. telescopes focus Cherenkov light on the camera by reflecting it on 382 spherical, aluminium coated, glass mirrors, which are distributed over the spherical frame of the telescope dish. All mirrors are mounted on the dish via a set of actuators that allow for calibration of each mirror individually. Each mirror has a 30-m radius

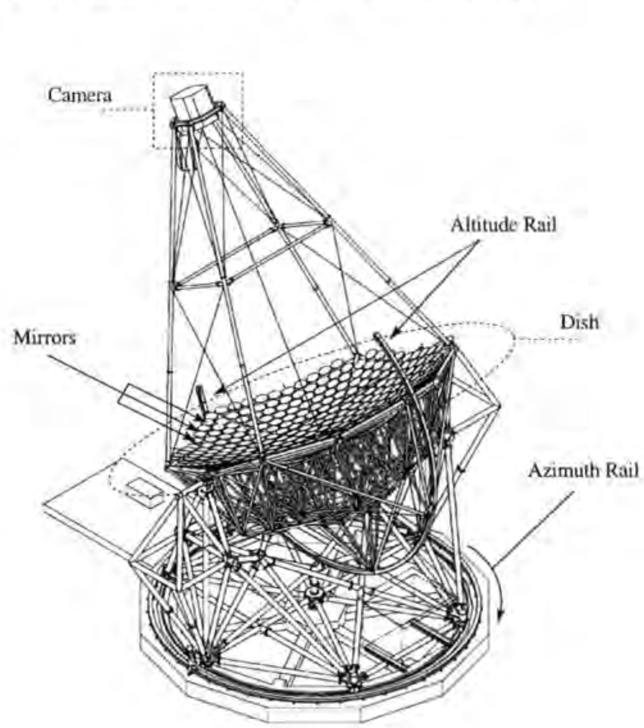


Figure 4.31: (from [331]) Technical drawing of the rear view of a H.E.S.S. telescope.

of curvature and a focal length of $f = R/2 = 15$ m. The flat-to-flat profile of the mirrors is 60 cm, and the overall reflective area per telescope is 108 m^2. A quartz overcoat applied on top of the aluminium layer prevents it from being damaged by dust particles, etc., without blocking significant amounts of light. The total reflectivity of the mirrors in the 300–600 nm range is $> 80\%$

The Davies–Cotton Design

It was mentioned that the H.E.S.S. telescopes are segmented reflectors. The specific design according to which the individual segments are placed and orientated on the dish frame is the well-known Davies–Cotton design [342]. According to the design, the spherical mirror segments are mounted on a spherical frame whose radius of curvature, R, is half that of the individual segments. Furthermore, each mirror's axis is orientated to converge at a distance $2R$ along the telescope axis, as measured from the *vertex* of the spherical dish: i.e. the point where the telescope axis intersects the spherical surface (see Fig. 4.33). In general, the choice of such a design should serve two main purposes:

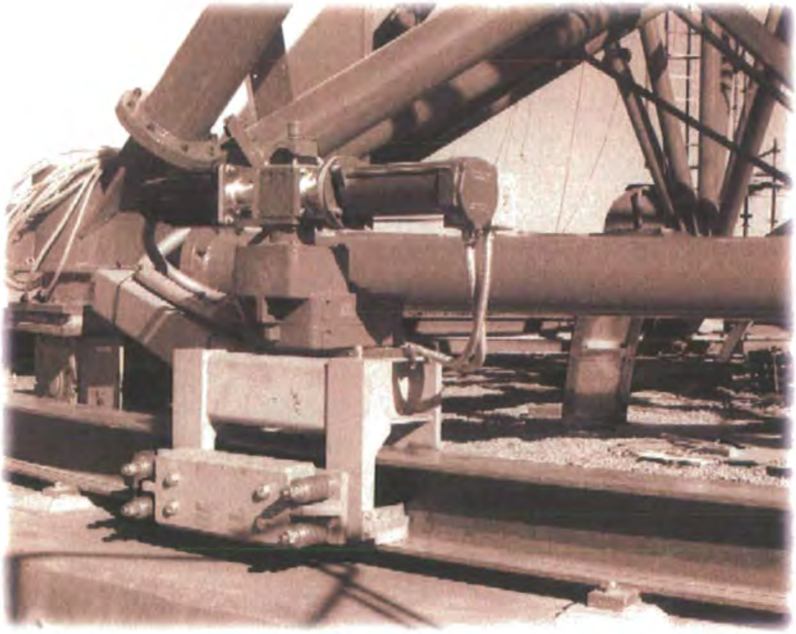


Figure 4.32: (from [331]) A close-up of the azimuth drive motor of the H.E.S.S. telescopes.

- **Good on/off-axis performance**. Cherenkov telescopes can observe across a fairly wide FoV ($> 3°$). In order to keep the image quality free of aberrations across the whole FoV, the detector's optics should be able to accurately resolve on-axis from off-axis events. In other words, the projected images on the camera plane should be as close to reality as possible and, thus, reflect the true angular width and position of the source.

- **Small spread of optical path delays**. It is also essential to retain the temporal spread of the Cherenkov showers upon reflection on the various mirrors across the dish. The time delays introduced by the different optical paths of the incident wave-fronts extend the duration of the Cherenkov flashes by an amount which depends on the reflector's design. This means that the required time window during which the PMTs have to register a signal excess (typically a few ph.e.) in order to trigger the telescope becomes wider. The recorded NSB during that time is also increased. As a result, the signal-to-noise ratio and the sensitivity to low-energy events decreases, which consequently increases the overall energy threshold of the detector. Eq. 4.15 shows how the latter is related to those parameters.

The first of these requirements is well met with a Davies–Cotton design. Normally, a single spherical reflector of radius R would display a large spherical aberration around the focus, which is caused by the outlying reflected beams. Only paraxial beams are focused correctly at a single point (Fig. 4.34). The Davies–Cotton design corrects these aberrations by using spherical mirrors tilted in such way as to be able to focus the light on a single point; *that* is where the camera centre is positioned: i.e. at the geometrical point from which all the mirrors are distance R away. Of course, each individual mirror also exhibits spherical aberration. However, because of the long radii of curvature compared to the mirror dimensions, all the reflected beams follow roughly the same path towards the focus of the telescope (see Fig. 4.34).

One important property of the Davies–Cotton design is the good off-axis performance compared to paraboloidal reflectors. The latter ones have increasingly detrimental coma aberration for off-axis incidence, whereas the overall spherical shape of the Davies–Cotton design reduces the global coma aberration [343],[341]. Consequently, Davies–Cotton reflectors have better concentration of light in the image. Fig. 4.33 shows schematically how the coma aberration is generated with multi-faceted reflectors.

Assuming that light from an extended source of angular width θ is reflected on a telescope's mirrors, the beams that arrive parallel to the telescope axis are all focused on the centre of the camera plane. Furthermore, the beams from the edge of the source form an angle θ with the axis, which causes them to land on a different point on the camera: in this case, to the left of the focal point. The angular width of the source, as it appears on the camera plane, relates to that of the actual source via the angle under which each mirror is viewed from the camera. It turns out that the width of the source's reflection on a mirror, viewed from the camera under an angle δ, is magnified with respect to the actual one by $1/\cos\delta$ times. The superposition of the images from all the mirrors produces a smeared image whose RMS width is proportional to

$$\sqrt{\langle (width)^2 \rangle} \propto \left(\frac{d}{2f}\right)^2 \theta \tag{4.24}$$

where d is the diameter of the dish's flat-to-flat profile [344].

As far as the second requirement for the reflector's design is concerned, spherical reflectors, in general, have a disadvantage compared to paraboloidal ones in that they are asynchronous surfaces. In other words, the different parts of a Cherenkov wavefront will not reach the camera of the telescope at the same time but will display a temporal spread due to the different geometrical paths followed by beams striking the different parts of the dish. A typical delay between light arriving from the edge and that from the vertex of the dish amounts to a few nanoseconds. For a 10-m spherical reflector, like those of H.E.S.S. and Whipple, the estimated delay is ≈ 6 ns [343].

In contrast, paraboloidal mirrors do not suffer from temporal scattering of the Cherenkov light, since they are isochronous surfaces. However, their construction is more expensive and their use in Cherenkov telescopes must consider off-axis performance as well. Fortunately, the Davies–Cotton timing can be improved. The use of small mirrors with higher f-numbers seems to present a solution for reducing the temporal spread. The f-number of the H.E.S.S. mirrors is $f/1.2$, which is substantially larger compared to that of previous Davies–Cotton designs (e.g. Whipple has $f/0.7$). This reduces the temporal spread by more that 2 ns, which makes it comparable to a Cherenkov event's [345].

In conclusion, the currently used Davies–Cotton design is a combination of good on/off-axis performance and low construction costs. Compared to normal parabolic

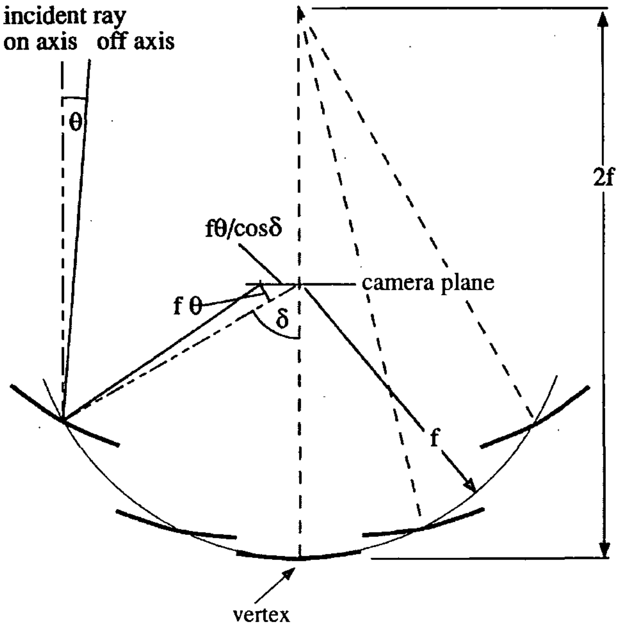


Figure 4.33: (from [344]) The basic geometry of the Davies–Cotton design. The individual mirrors of the telescope have focal lengths, f, equal to the radius of curvature of the dish frame on which they are laid. Each mirror's axis is adjusted to cross the telescope's axis at distance equal to $2f$. This way, the incident light beams that arrive at the dish, parallel to the telescope axis (dashed lines), will cross the centre of the camera which is situated at the focal point of the mirrors. On the other hand, the impact point of off-axis beams (solid lines) on the camera plane is shifted relative to the centre of the camera. This expresses the coma aberration, which is inherent in all spherical mirrors. Finally, the angular width, θ, of extended sources is lengthened by a factor $1/\cos\delta$, where δ is the angle under which the mirrors are viewed from the camera. Hence, off-axis performance deteriorates due to image smearing that is caused by the outlying mirrors.

reflectors, this design offers better off-axis performance, but introduces a small time-spread due to its asynchronous spherical surface. Furthermore, compared to a single reflecting surface, the tessellated configuration is easier to maintain, and the individual mirrors are much cheaper to built and replace.

More details on the specific Davies–Cotton layout of the H.E.S.S. telescopes can be found in [344].

Mirror Alignment

In practice, the 382 mirrors of the tessellated H.E.S.S. reflectors need frequent calibration in order to make sure that they all point in the right direction. The calibration is carried out first by pointing the telescope at a reference star. As discussed in section 4.3.1, this is done primarily with computer controlled drives on the Alt–Az axes but is fine-tuned with a finder scope, codenamed Sky CCD, which is mounted off the dish centre, in order to have a clear view of the sky. A different CCD scope, called Lid CCD, is mounted at the vertex of the dish and has a clear view of the camera plane. The Lid CCD has a substantially larger FoV (f = 180 mm) than the Sky CCD (f = 800 mm), as well as a much shorter depth-of-field with f/2.8: these properties have been chosen for the optimal coverage of the camera. The mirror calibration takes place with the camera CCD closed. Then, all the individual reflections of the reference star on the camera lid produce a pattern of well-defined spots (see Fig. 4.35), which is shown in Fig. 4.36a. By adjusting each mirror's orientation, all the spots can be made to converge on the camera centre (see Fig. 4.36b). Consequently, all photons arriving parallel to the telescope's axis will be focused on the centre of the camera. The accuracy of the whole calibration process is limited by the resolution of the Lid CCD.

Finally, the Lid CCD is used for monitoring the telescopes' mast deformations. The weight of the masts can cause them to bend at different degrees depending on the different altitude positions of the telescope. For that reason, the FoV of the Lid CCD is slightly larger than the camera's size, in order to include 3 LEDs positioned around the camera. This LEDs help as reference points for monitoring the degree of mast deformation. Using that information, the mirror calibration can be adjusted accordingly [46].

The whole process of mirror alignment is automatic and regularly performed in order to ensure undistorted projection of the Cherenkov images. The reader can find more about the process in [346].

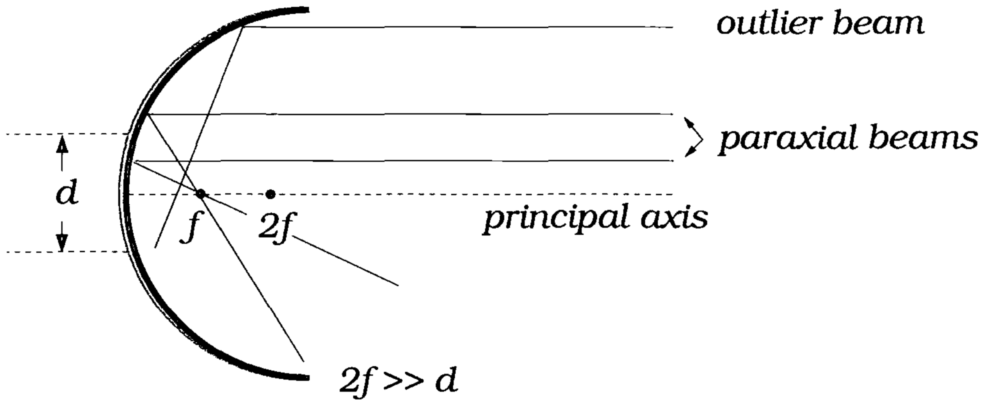


Figure 4.34: The spherical aberration of an individual mirror in the Davies–Cotton design is negligible because its radius of curvature, $R = 2f$, is much longer than the mirror's profile. Even the outermost beams are practically focused on the focal point f.

Camera

Each of the four H.E.S.S. telescopes has a camera placed at the focal plane of the telescope, which is located along the telescope's axis, $\approx$ 15 m away from the vertex of the spherical dish. The cameras have a roughly hexagonal shape which is $\approx$ 1.5 m in diameter. That provides them with a 5° FoV, which is large enough to cover the full apparent size of expected extended sources (see Fig. 4.25).

Each camera's surface is covered with 960 PMTs: their front windows are made of borosilicate glass, and their sensitivity is optimum between 300–600 nm, which covers the part of the Cherenkov spectrum where most of the light is emitted (see Fig. 4.37). The PMTs themselves have very narrow viewing angles (0.16°). However, all PMTs have Winston cones fitted on them, which allows the coverage of the in-between empty space with their reflective surfaces. This way, all incident light beams are redirected into the active cross section of the PMTs. The extended viewing angle, with the use of Winston cones, reaches $\approx$ 30°, which is the optimum cut-off angle for the detection of all the reflected light from the dish, but it also helps to reduce the ambient light coming from wider angles outside the dish [347].

The maximum trigger rate that the camera can process is $\sim$ 1 kHz. All the required electronics for image digitisation, triggering and readout of the recorded

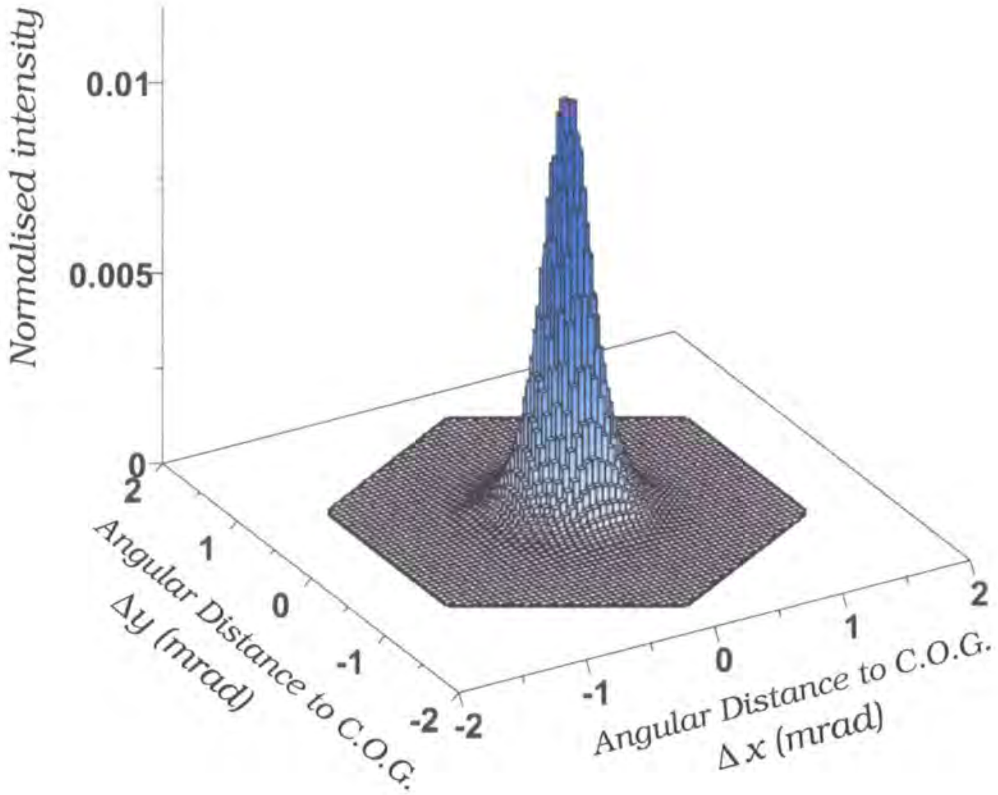


Figure 4.35: (from [337]) The intensity distribution of a point light source, otherwise known as Point Spread Function (PSF), as recorded by the Lid CCD which is mounted on the vertex of the telescope dish. The hexagonal frame on the x–y plane corresponds to the size of a PMT on the camera, and the distance from its Centre of Gravity (C.O.G.) is measured in mrad. Finally, the vertical axis represents the light intensity in arbitrary, normalised units.

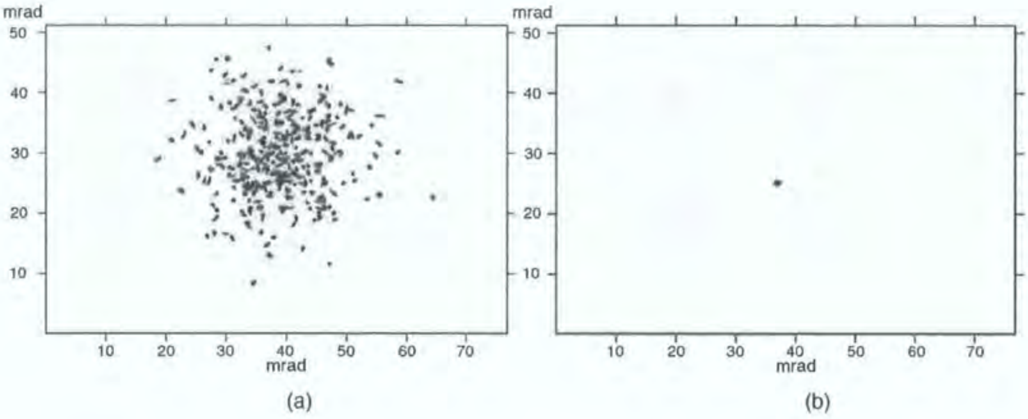


Figure 4.36: (from [331]) The 382 reflected images of the reference star, as viewed with the Lid CCD, (a) before and (b) after calibration of the H.E.S.S. mirrors. The coma aberration which distorts the images coming from the outlying mirrors is evident.

events are integrated into the camera body. The trigger criteria are customised to minimise accidental NSB triggers. Their application concerns the different sectors in which the camera is divided, and each sector consists of 64 contiguous pixels. Furthermore, to ensure a homogeneous trigger efficiency, these sectors have overlapping regions. In each sector, a trigger is registered if at least p pixels (sector threshold) receive a minimum of q ph.e. each (pixel threshold). All the triggered pixels in a sector have to report a signal within a time window which is typically $\tau_{\mathrm{w}} \approx 16$ ns. Such a short interval guarantees the suppression of NSB triggers, camera-wide. System-wide, the array is rarely triggered by coincident NSB events. However, the rejection of NSB events by the camera happens in a much shorter time (≈ 5.5 μs) than that by the whole array (≈ 450 μs). Hence, NSB events are only responsible for increasing the dead time of the camera trigger: i.e. the time delay due to processing (readout or rejection) of events before sampling of new events can take place [335].

Clearly, the selection of an appropriate trigger criterion is essential as it affects both the trigger rate and the energy threshold of the H.E.S.S. array. The crucial parameters are q, p and τ_{w}. A very loose criterion, i.e. low q and p, would increase the trigger rates and lower the energy threshold of the system, since it would allow low-ph.e. events to trigger the system. The trigger rate is also expected to be highly unstable and, as a result, quite unmanageable. On the other hand, a strict criterion

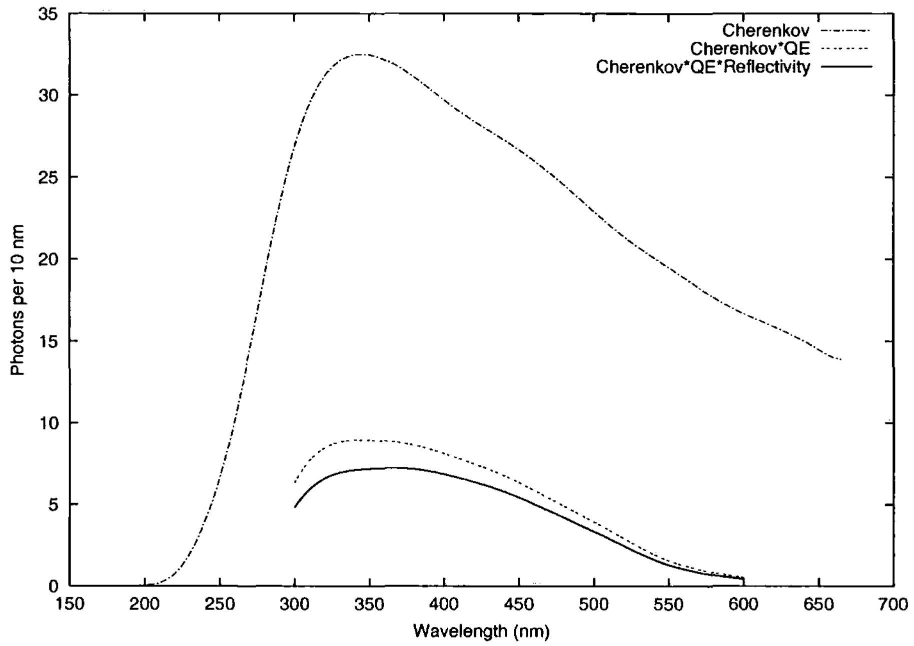


Figure 4.37: (from [349]) The simulated differential photon spectrum of the Cherenkov emission for 1-TeV γ-ray showers, as detected from the H.E.S.S. site (dashed-dotted line). The dotted line represents the resulting spectrum with the PMT quantum efficiency (QE) included, and the solid line represents the same spectrum after including the reflectivity of the mirrors as well.

would discard many low-energy events, which would increase the energy threshold of the system. Fig. 4.39 shows a plot of the trigger rate's dependence on the pixel threshold for single and stereo observations. The data were collected with all 4 H.E.S.S. telescopes pointing at a relatively dark part of the sky close to zenith [335]. The spectral break in the event rate distribution is evident. Below a certain pixel threshold, the trigger rate is dominated by the soft NSB spectrum, whereas at higher event energies the air shower triggers are more dominant (see section 4.2).

Calibration and monitoring of both the individual PMTs and the camera as a whole is carried out by purpose-made LEDs. Each PMT's operation is checked frequently with dedicated LED systems that are attached to the inside of the camera lid (see Fig. 4.38). When the telescope is not observing, the camera lid is closed to allow the LEDs to create 1-ph.e., artificial signals, in order to check the individual ph.e.-to-d.c. conversion coefficients [348]. The next paragraph explains the details of the calibration of the camera as a whole, which is called *flat-fielding*.

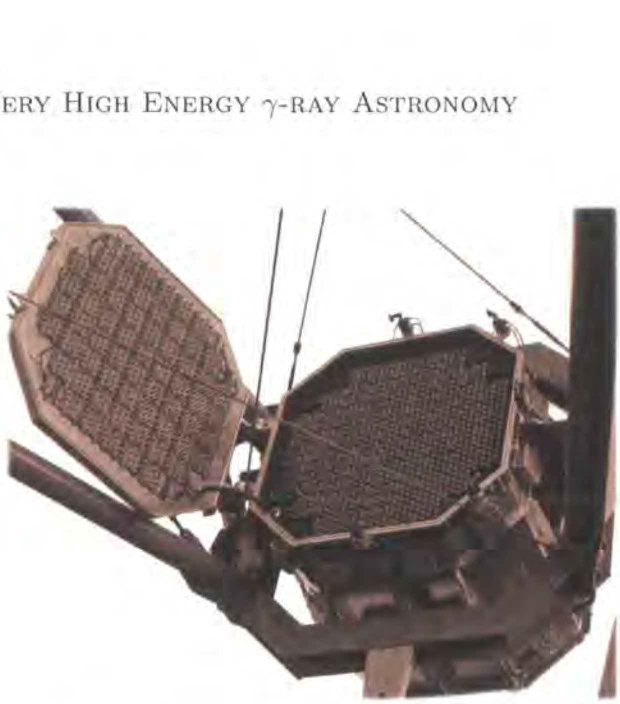


Figure 4.38: (from [331]) The H.E.S.S. camera with its lid open. The circuit plate on the visible side of the lid carries LED systems — one per PMT — that generate 1-ph.e. pulses, which help determine the individual PMT response, when the telescope is not observing.

Flat-Fielding

Knowledge of the PMT response prior to observations is essential in understanding the magnitude of the observed signals. For that purpose, the camera is regularly subjected to artificial Cherenkov-like flashes that are generated from an LED system situated near the centre of the telescope dish. The purpose of this calibration process, called flat-fielding, is to measure the gain variations of the PMTs across the camera. By doing that, we are able to translate correctly the amount of d.c., which the ADC unit has registered in a PMT, to an amount of ph.e. and, consequently, to photon flux. Of course, when flat-fielding, the artificial source has to produce a known amount of light, so that combining this information together with the amount of d.c. leaves as the only unknown the gain of each PMT. For that reason, each reference flash has to have uniform intensity across the camera in order to reveal the true variations in the gains of the different PMTs. When flat-fielding, the condition for triggering is made more strict, requiring > 9 neighbouring pixels to be illuminated at the same time to register an event. In that way, the number of background, true Cherenkov events is reduced [324].

The flat-fielding device consists of three pulser circuits, designed and built at

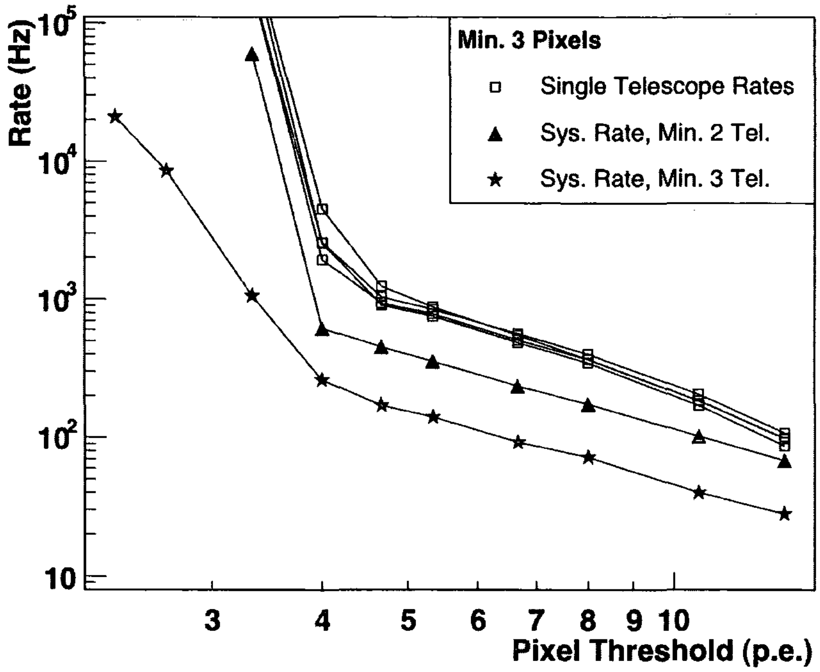


Figure 4.39: (from [335]) Dependence of the trigger rate (events s^{-1}) on the pixel threshold q, for single (clear squares) and stereo observations. The solid triangles and stars correspond to the requirement of a minimum of 2 and 3 telescope triggers, for a system trigger, respectively. The minimum sector threshold was set to 3 pixels and the error bars are smaller than the symbols. Regardless of the number of telescopes, the rapid increase of the trigger rate as the trigger criterion becomes loser is evident. Compare this figure with Fig. 4.9 which explains the two components of the event-rate distribution.

Sheffield University, that are capable of driving 3 LEDs to simultaneous flashes of $\approx$ 5 ns. The temporal width of each flash was defined as the time interval between the moments at which the radiation intensity of the flash is half that of the maximum: i.e. equal to the Full Width at Half-Maximum or FWHM. This pulse width is a good match to the real Cherenkov events, which means that no special adjustments in the software or hardware are needed to record the flat-fielding events. The choice of flat-fielding LEDs for H.E.S.S. was based both on the close match to the peak of the Cherenkov spectrum, but also on factors like availability, brightness and cost. Although it is possible to use a combination of LEDs that would cover a larger part of the Cherenkov spectrum, it was decided — for the initial operation of the device — to use 3 near-UV LEDs from HERO Electronics [350]. The LED spectrum covers the wavelengths from 390 to $\approx$ 410 nm, with a half-intensity angle of 20°. Considering the $\approx$ 6° angular width of the camera as seen from the flat-fielding device, the beam width is more than enough to contain the whole camera in its centre, where the intensity varies the least. However, it was also considered appropriate to use a diffuser to ensure that any possible anisotropies across that area are eliminated. Finally, a filter wheel with 5 neutral-density filters was placed in front of the 3 LEDs and behind the diffuser. These filters have different grades so as to control the intensity of the flashes at the camera. It is a well-known effect that a PMT's response varies with the intensity of the incident light. Hence, a more accurate calibration is achieved by checking the variation of the responses across the camera for a wide range of intensities.

The flat-fielding devices on all 4 H.E.S.S. telescopes perform well. Prior to installing the system, the device was checked for intensity variations between subsequent pulses: it was found that the pulse intensities vary by $< 5\%$ RMS. After installation, the first tests showed that the whole camera is illuminated adequately and uniformly by flashes whose flux ranges from 2.5 to 250 ph cm^{-2}, at the camera, depending on the filter-wheel setting. However, it is necessary to monitor the performance of the device, as it might degrade after a long period of operation. For that reason, there is an additional photo-diode installed between the filter wheel and the diffuser, which provides the means for quick and easy-to-use monitoring of the flash intensity and pulse characteristics. A schematic of the flat-fielding circuit is presented in Fig. 4.40.

Apart from producing stable and uniform flashes, the flat-fielding device is also capable of pulse repetition rates of up to 1 kHz. The repetition rate can be con-

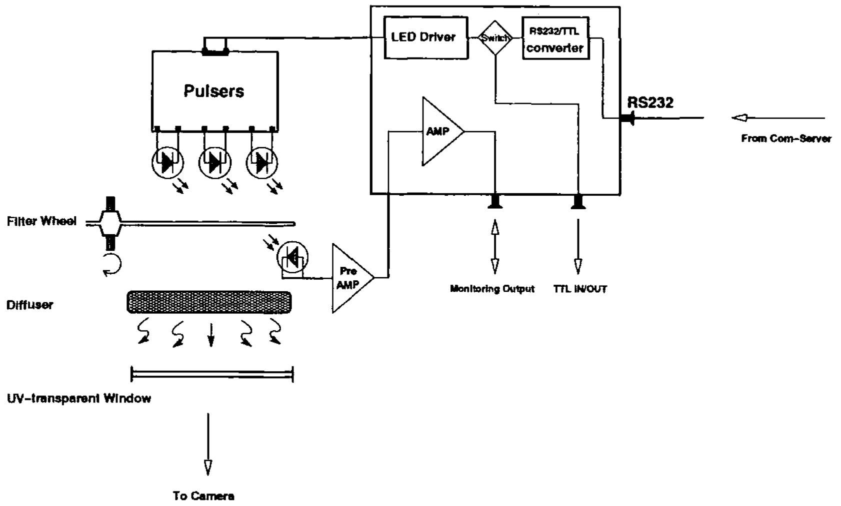


Figure 4.40: (from [349]) The design of the flat-fielding device used in the H.E.S.S. telescopes. The frequency of the LED pulses is controlled either remotely, with a com-server, or locally, with a TTL pulse generator. Using the appropriate driver, the TTL signal is converted to a compatible-with-a-Sheffield-LED-pulser pulsed signal. There are 3 pulsers integrated into the flat-fielding circuit. Each pulser carries a near-UV LED which is made to flash simultaneously with the rest. The intensity of the flashes is controlled by placing a filter wheel in front of the LEDs, which uses 5 neutral-density filters (grades 0.5, 1.0 ,1.3, 1.5 and 2.0) to attenuate the beam to the preferred intensity. The forward beam is then diffused by a holographic diffuser, in order to eliminate any intensity variations across the beam's cross section. Finally, the apparatus is shielded from the outside with a UV-transparent window placed in front of the diffuser. An additional monitoring photodiode is placed in front of the filter wheel, providing a quick and easy check on the pulse characteristics that are expected to deteriorate with time.

trolled locally or remotely via a pulse generator or a com-server trigger, respectively. Although such frequent triggering is not directly useful for the calibration process, it is a good way of checking the response of the Data Acquisition System (DAQ) under high incident event rates.

Table 4.1 lists the basic characteristics of the pulses generated by the flat-fielding device in its current configuration, which uses 3 HUVL400-520 LEDs by HERO Electronics. The performance of the device is described in the conference publication [349].

Wavelength	390–410 nm
Pulse width (FWHM)	5 ns
Pulse rise time	2.5 ns
Short-term temporal stability (jitter)	< 0.5 ns
Long-term temporal stability (drift rate)	< 0.25 ns y^{-1}
Intensity stability	$< 5\%$ RMS
Photon flux per pulse at 15 m	250 ph cm^{-2}
Max. trigger rate	1 kHz

Table 4.1: (from [349]) The pulse characteristics of the H.E.S.S. flat-fielding device. The LEDs used in this configuration are the HUVL400–520 by HERO electronics.

Central Trigger (CTS)

The trigger decisions for the full array of H.E.S.S. are made by the Central Trigger System (CTS) which is located at the control building of the array. Each telescope reports a local trigger to the CTS, and the latter makes a decision for a system trigger, based on the time difference between the different local triggers (see Fig. 4.41). At least two telescopes have to report a local trigger to the CTS within a few μs, in order to form a system trigger. Otherwise, the local events are discarded, and a reset signal is sent to each camera so that further triggering can progress. The telescope distances and pointing directions are all accounted for in the timing-coincidence calculation. To minimise the round-trip time of the information between each camera and the CTS building, the connections are carried out through optical fibres; this results in a 4.2-μs average communication delay.

As Fig. 4.39 shows, for a minimum of 3 triggered telescopes, as a system trigger requirement, the trigger rate is reduced as much as an order of magnitude with respect to that from single-telescope observations. As a result, the pixel threshold can be further reduced without the risk of increasing the event-read-out dead times beyond the Data Acquisition System's limitations. Hence, the array can operate at lower thresholds than a single telescope. The role of the Data Acquisition System (DAQ) is explained in the next paragraph.

Unfortunately, a multi-telescope trigger requirement will discard showers which are not energetic enough to trigger the minimum required number of telescopes. Hence, although stereoscopic observations can reliably characterise air showers down to lower energies than single-telescope observations, this is potentially done on the expense of pulsed, low-energy events.

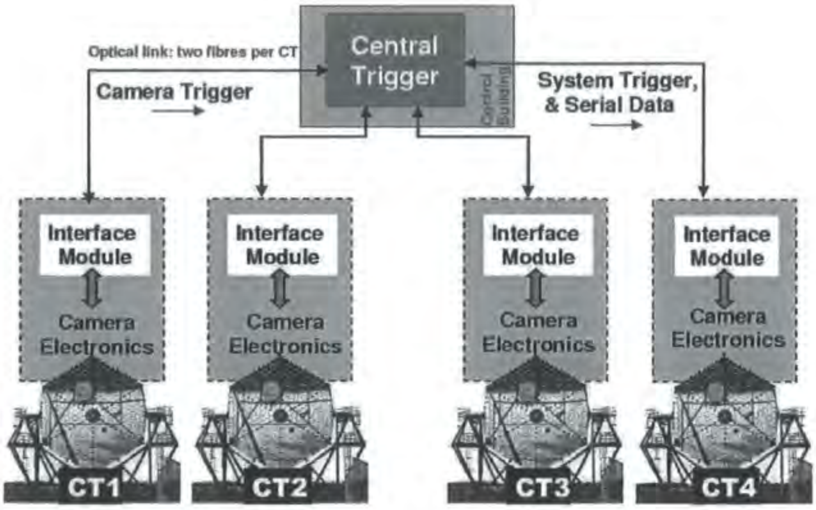


Figure 4.41: (from [335]) The control flow of the H.E.S.S. Central Trigger System (CTS). Each telescope reports a trigger to the CTS, where the decision for a system trigger is made. The communication between the telescopes and the CTS is carried out via optical fibres which connect each telescope's camera interface with the CTS; the latter is located at the control building.

A complete description of the H.E.S.S. central trigger system can be found in [335].

Data Acquisition System (DAQ)

After a trigger decision has been taken, the data is read out and stored on large hard discs on-site. Each event occupies ≈ 1.5 kB of memory, and hence, at a maximum trigger rate of 1 kHz, four telescopes would produce ≈ 6 MB s^{-1} of data or ≈ 100 GB per observation night. Clearly, a system that manages all this information in a stable manner is required. This system is the Data Acquisition System (DAQ), whose responsibilities are the following:

- **Read-out and storage** of events generated from the camera, tracking, pointing and atmospheric monitoring sub-systems. By far, the largest amount of data is generated from the camera, whereas the size of data from other subsystems is small [351].

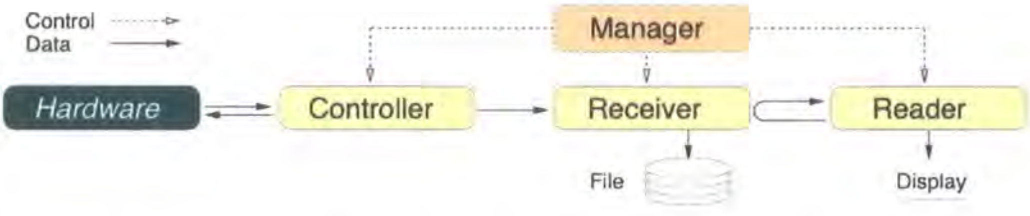


Figure 4.42: (from [351]) The processes of the H.E.S.S. Data Acquisition System (DAQ).

- **Synchronisation and monitoring** of all the above sub-systems, as well as error-handling during those processes.

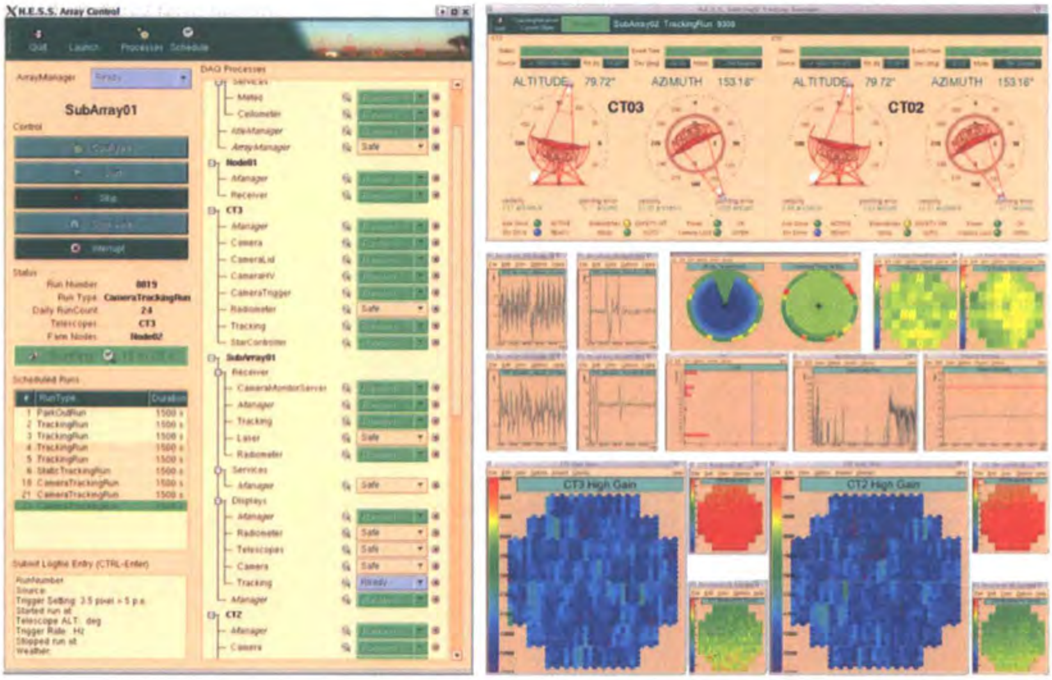
In general, the DAQ of H.E.S.S. Phase I is based on a Linux computer cluster, which processes and transports the data using C++- and Python-based interfaces. Data from each observation run is stored as binary files, with the use of the ROOT Data Analysis Framework. The framework also provides graphical routines for an intuitive visualisation of various parameters, like trigger rate, weather data, run-by-run statistics, etc.

More specifically, the DAQ framework is composed of 4 distinct *processes* (see Fig. 4.42): the Controllers, which interact directly with the hardware and read out the data; the Receivers, which process and store the data in the appropriate format on disks; and the Readers, which request part of the data asynchronously. The purpose of sampling data with the Readers is so that they can be visualised for the shift crew. An example of such visualisation is presented in Fig. 4.43, where one can see the central Graphical User Interface (GUI) and some monitoring displays.

Finally the Manager process, as its name suggests, controls the data-taking by interacting with the rest of the processes. For example, one of the Manager tasks is to read the observation schedule and assign a unique run number to each sequence (e.g. #16,000). An in-depth presentation of these processes and the H.E.S.S. DAQ, in general, is given in [351].

Weather Station

Weather monitoring is an integrated part of the H.E.S.S. DAQ. A weather station that operates continuously is installed at the H.E.S.S. site and is capable of taking measurements of the air temperature, relative humidity, atmospheric pressure, wind



Status

Run Number 8819

Run Type CameraTrackingRun

Daily RunCount 24

Telescopes CT3

Farm Nodes Node02

Scheduled Runs

#	RunType	Duration
1	ParkOutRun	1500 s
2	TrackingRun	1500 s
3	TrackingRun	1500 s
4	TrackingRun	1500 s
5	TrackingRun	1500 s
6	StaticTrackingRun	1500 s
18	CameraTrackingRun	1500 s
21	CameraTrackingRun	1500 s

Submit Logfile Entry (CTRL-Enter)

RunNumber
Source
Trigger Setting: 3 5 pixel > 5 p.e.
Started run at
Telescope ALT: deg
Trigger Rate: Hz
Stopped run at
Weather

Figure 4.43: (from [351]) The central Graphical User Interface (GUI) of the H.E.S.S. array (left), and some examples of monitoring displays (right).

Figure 4.44: (from [351]) The weather station.

speed, wind direction and rainfall (see Fig. 4.44). These parameters are checked in the post-processing of camera data in order to assess the quality of the observations. Furthermore, weather station data can be used in atmospheric modelling, with software packages like MODTRAN (MODerate resolution TRANsmittance code).

Ceilometer

The H.E.S.S. ceilometer is a LIDAR (LIght Detection And Ranging) that emits a $\approx$ 900-nm laser beam vertically into the sky and detects the amount of backscattered light — mainly due to aerosols — out to 7.5 km. It is also capable of measuring the overhead altitude of cloud layers. The instrument reports the amount of backscattering in units of $\mathrm{km}^{-1}\ \mathrm{sr}^{-1}$, which can be inverted to recreate the optical profile of the atmosphere. By comparing the resulting profile with model profiles of MODTRAN, one can calculate the extinction in the wavelength range of interest (250–700 nm) [352],[353]. A photograph of the installed LIDAR is shown in Fig. 4.45.

Infrared Radiometer

The infrared radiometer measures the temperature of the sky by comparing the incident IR emission within its 2.9° FoV with that of a blackbody spectrum, and

Figure 4.45: (from [354]) The H.E.S.S. LIDAR.

deduces the radiative temperature of the sky. Its sensitivity range is 8–14 μm, in which the atmosphere is transparent to the water vapour continuum emission (see Fig. 4.46). Hence the device is very sensitive to the presence of clouds, which appear warmer compared to the rest of the sky, due to their higher water content. Furthermore, changes in relative humidity and temperature can still cause night-to-night variations of the sky temperature, regardless of the presence of clouds. In both cases, there is a clear correlation between sky temperature and trigger rate during an observation run. As Fig. 4.47 shows, the trigger rate during the observation run 9888 was increased by $\approx 20\%$ after the overhead clouds cleared out.

Although the appearance of overhead clouds during observations cannot be avoided, monitoring the sky with the infrared radiometer can help us discard the offending runs. Such runs are very likely to contain Cherenkov events whose light was heavily attenuated in their journey to the telescopes, and therefore their observed intensities are misleading.

Each of the four H.E.S.S. telescopes is equipped with a paraxially installed radiometer, thus providing the means for sky monitoring in the direction of observation (Fig. 4.48a). In addition, a scanning radiometer operates on-site, which allows us to have an overview of the whole overhead sky and warns us of any approaching

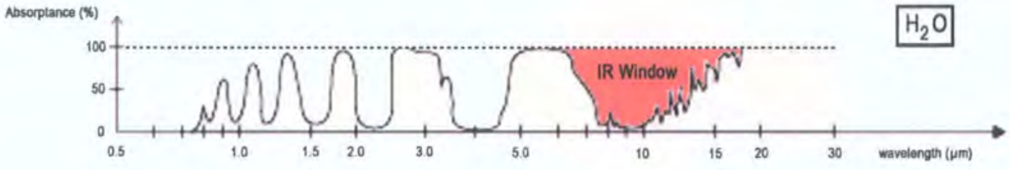


Figure 4.46: (from [355]) The absorptance spectrum of the water-vapour continuum emission. The infrared radiometer is sensitive to IR emission in the range 8–14 μm, across which the atmosphere is transparent to the thermal emission of water vapours.

weather fronts (Fig. 4.48b).

4.3.2 Supporting Equipment

H.E.S.S. Timing

A supporting device which can provide a direct verification of the correct function of the H.E.S.S. timing system is installed on the camera lid of the first H.E.S.S. telescope (CT3). The device is designed to detect optical pulsed emission by collecting light after it has been reflected on the telescope dish. The Custom-Built Detector (CBD) is located at the focal plane of the telescope. Its main components are a PMT which registers the pulsed emission, a secondary plane mirror which redirects the light from the focal plane of the telescope directly into the PMT and, finally, an integrated system of fast electronics for the digitisation and the time assignment of events. The signal from the PMT is digitised with a 16-bit ADC that samples at 20 kHz. The timestamp assignment for each event is derived from the GPS clock of H.E.S.S.'s CTS, which has an accuracy of < 1 μs.

Clearly, the applicability of the CBD is towards pulsar observations. In order to study the γ-ray emission of millisecond pulsars, one needs to keep track of the pulsar's phase during observations across the year: the diurnal and annual motion of the Earth are the main reasons for the variability of a pulsar's period, as seen from the observatory; in addition, the period changes monotonically with time because of the intrinsic pulsar spin-down. For pulsar observations below the detector's energy threshold, one must rely on the periodicity of the signal to achieve a detection. Therefore, an accurate ephemeris is invaluable for such observations. However, it is often true that such ephemeris is not available for all observing times, and, moreover,

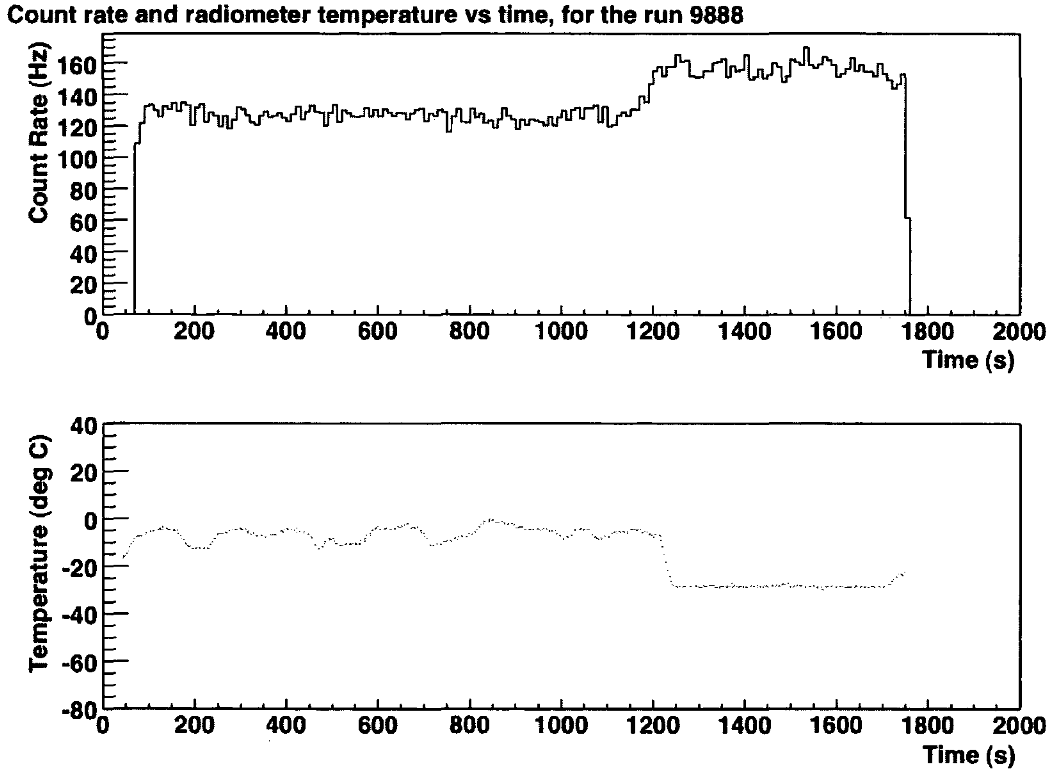


Figure 4.47: (from [352]) Correlation between the sky temperature in the direction of observation and the trigger rate. The presence of clouds in the first 1,200 seconds of run 9888 is evident by the sudden drop in the sky temperature, as was recorded with the paraxial radiometer, immediately afterwards. This caused the trigger rate to rise by $\approx 20\%$ for the last 600 seconds of the run.

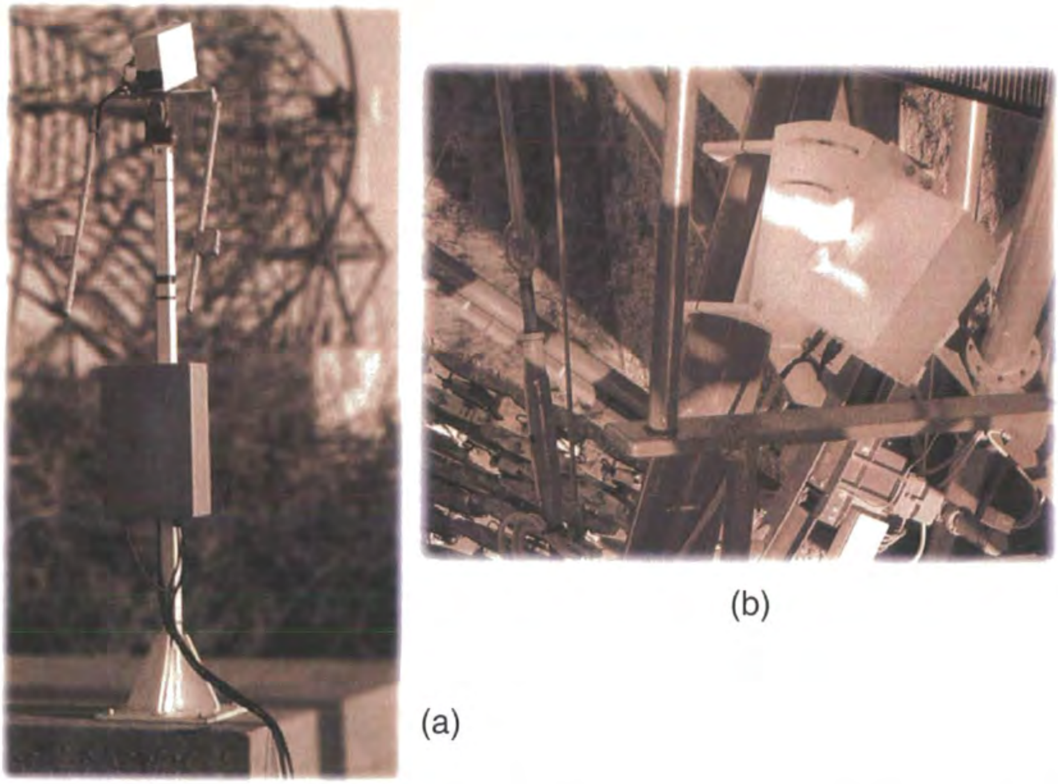


Figure 4.48: (from [356]) The scanning (a) and paraxial (b) infrared radiometers at the H.E.S.S. site.

pulsar glitches can cause even a recent ephemeris to become invalid. Hence, there is a clear motivation for having the CBD together with H.E.S.S., since it can provide a cross-check between the various ephemerides and the actual observations; but also, it can verify the barycentring routines that are used to transform the pulse arrival times from the observatory to the Solar System Barycentre (see section 5).

The CBD has been tested with the Crab and Vela pulsars, from which 6 and 10 h of good quality data were collected, respectively. Data that were affected by the bad weather conditions were discarded. The timestamps were folded into phases using the Jodrell Bank ephemeris [357], and then they were sliced into 10-s intervals, whose data quality was checked [358]. Only the slices whose background levels were below an acceptable threshold were considered in the analysis; and approximately 3 h of data were selected for the Crab pulsar. This background was then subtracted from each event's *size* and the resulting excess/deficit was binned according to each event's corresponding phase. The lightcurve shown in Fig. 4.49 comes from the summation and binning of all the d.c. contributions from all the analysed slices. This lightcurve is consistent within experimental errors [358] with optical data from the Crab pulsar taken by other experimenters [359].

Finally, the Vela pulsar data were also analysed in the same way, but no significant pulsation was observed, and the estimated exposure time for a 5-σ detection was set to 30 h. Future versions of the device will include more efficient, multiple light sensors for the exclusion of optical transients (e.g. meteors).

The above process of extracting the Crab pulsar's lightcurve from optical data is presented in more detail in section 6.3.1, where the function of a custom code written for this purpose is presented.

Lights on Hills

An additional atmospheric-monitoring instrument which provides direct measurements of the local atmospheric extinction, as well as the means for a better interpretation of the LIDAR measurements, has been operating at the H.E.S.S. site. The instrument is a *transmissometer* that assesses the atmospheric transmissivity across a horizontal distance of 30 km and a vertical of $\approx$ 550 m. This is done by measuring the attenuation that a reference light has sustained after crossing the above distance through the atmosphere.

The location of the light source is at the top of the Gamsberg plateau, which is $\approx$ 30 km away, at an altitude of 2350 m a.s.l. This natural land formation offers

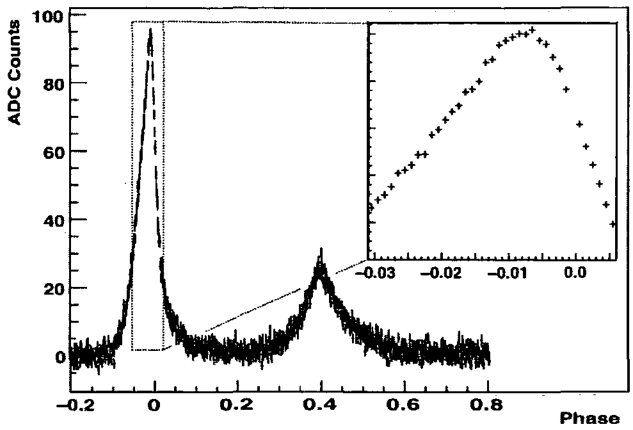


Figure 4.49: (from [358]) The Crab pulsar's optical lightcurve. The lightcurve was constructed from 3 h of data, which were taken with the Custom-Built Detector mounted on the camera lid of CT3. The position of the main peak could be determined with an accuracy of $\approx 13\ \mu$s.

excellent optical clearance because of its elevation with respect to the telescope's site. Moreover, it is close enough to be clearly visible but also far enough to provide a useful degree of attenuation.

The reference light source can be adjusted to emit at four different wavelengths, 390, 455, 505 and 910 nm. The first three wavelengths were specifically chosen for the measurement of the transmissivity at Cherenkov wavelengths. In addition, the 910-nm light matches roughly the wavelength of the LIDAR beam (optimised for the water vapour transmission window) and it will be used for cross-calibration between the LIDAR and the transmissometer. Furthermore, the apparatus is equipped with monitoring devices that check the performance and operation of the light source and the rest of the components (e.g. intensity and battery power).

The intensity of the attenuated light, as it reaches the H.E.S.S. site, is measured with an appropriate detector. A rough schematic of the on-site transmissometer's set-up is presented in Fig. 4.50.

Amongst the planned experiments with this device is the cross-calibration between the ceilometer and the transmissometer. This is done by measuring the backscatter at 905 nm and the transmissivity at 910 nm in the direction of Gamsberg. Then, one could try to link the results and gain a better understanding of the LIDAR data. More specifically, it is possible to extrapolate the LIDAR measurements (905 nm) to the transmissometer's wavelength (910 nm), and deduce the

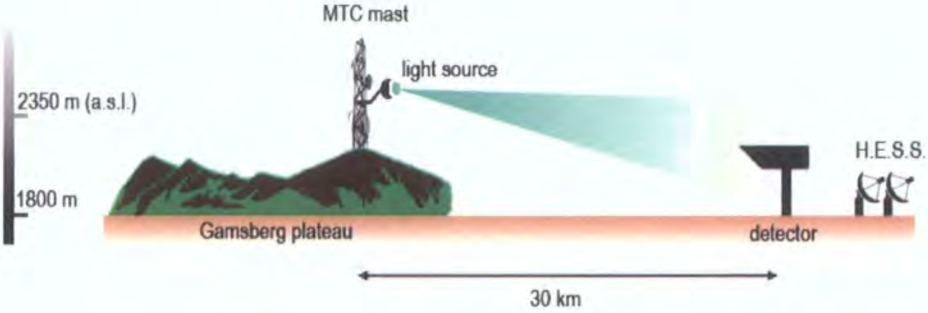


Figure 4.50: (from [360]) A schematic of the transmissometer set-up.

transmissivity of the atmosphere for altitudes up to 7 km above the H.E.S.S. site, in the 390–505 nm range. This can be achieved by using ceilometer data in combination with the ratios between the transmissivity in the 390–510 nm range and that at 910 nm, and between the backscatter and transmissivity in the direction of Gamsberg.

A brief description of the transmissometer device is given in [360]. The reader can find information on the principle of LIDAR–transmissometer intercomparisons in independent publications: e.g. [361],[362] and [363]. The idea of atmospheric monitoring at ground level is not new; other high-energy experiments have used similar devices, e.g. the Horizontal Attenuation Monitor (HAM) of the Pierre Auger fluorescence experiment (see [364]).

Optical Support

ROTSE

Apart from the optical instrumentation installed on the Cherenkov detectors, whose purpose is to align and calibrate the telescopes, a separate optical telescope is situated on the H.E.S.S. site. The ROTSE-3c is a multi-purpose, fully robotic equatorial telescope with fast slewing capabilities: both axes reach a maximum speed $\approx 35^\circ/\text{s}$. It has a Cassegrain focus with a 450-mm diameter, $f/1.8$ primary mirror. Its maximum FoV covers 2.64° in the sky, and it has been optimised for observations in the 400–900-nm band. A detailed description of the instrument exists in [365].

ROTSE-3c's purpose is to provide support for the high-energy observations by conducting the following optical observations:

- **Patrolling of the overhead sky and triggering on optical transient sources.** During normal operation, the ROTSE-3c telescope covers 100 3.5-deg^2 areas each night. This observation mode has been appropriately named sky patrol. In case of an unusually bright optical activity in ROTSE's FoV, the latter can immediately notify the H.E.S.S. telescopes, which can then conduct follow-up high-energy observations of the source in question.

- **Monitoring of potential γ-ray sources** that exhibit variability in the optical spectrum: e.g. AGN and microquasars. If a flaring state is detected at optical wavelengths, the H.E.S.S. crew can plan subsequent Cherenkov observations that could also reveal γ-ray activity.

- **Supporting observations at GRB alerts.** The fast slewing of ROTSE-3c will allow it to observe GRB events almost immediately after a notification. The collected information will be available for future observations, with the H.E.S.S. telescopes, of the refined GRB sample.

- **Multi-wavelength observations of H.E.S.S. targets.** ROTSE-3c is the only optical observatory on-site, and it is suitable for parallel observation of the scheduled H.E.S.S. sources in the optical band. Wide-band or multi-band photometry can provide valuable information on correlated activity between optical and γ-ray emission.

The ROTSE-3c instrument observes roughly 30 objects per night. Observations carried out from September 2003 to January 2004 amounted to 564 h of data plus 52208 images [366]. Amongst others, ROTSE's targets include the quasar PKS 2155−304, the BL Lac H 1426+428 and the HMXRB RX J1826.2−1450. A more inclusive list of targets, as well as the source-selection criteria, can be found in [365].

ATOM

A second optical telescope is in the progress of preparation and transportation to the H.E.S.S. site. It is based on a Zeiss, 20-year-old prototype that had been in the service of the Landessternwarte institute in Heidelberg until recently. The telescope is a Richey–Chretien design, which offers coma-free images across the whole FoV. It has a 75-cm hyperbolic primary mirror and an Alt-Az mount. After having received several upgrades, the telescope has been dismantled in order to be transported to Namibia.

Prior to installation at the site, the telescope will be equipped with mechanical and electronic units that will render it fully automatic. Furthermore, with the appropriate dedicated software, it will be possible to perform observations remotely and without the need for on-site personnel. In its final form, it will be appropriately named Automatic Telescope for Optical Monitoring (ATOM). Amongst ATOM's duties will be to monitor AGN variability, to measure the atmospheric extinction along the line-of-sight of the H.E.S.S. telescopes, and to carry out multiwavelength observations in parallel with the H.E.S.S. telescopes [367].

4.3.3 Observation Modes

The way observations are performed with the H.E.S.S. array has been evolving constantly. At the beginning of the experiment, when there was only one telescope, the single-telescope observations were divided into single ON/OFF and single wobble modes (see e.g. [156]). The former involves pointing directly at the source for the duration of a whole run and then, for an equal amount of time, sampling the background from a part of the sky which is ahead or behind the source by a certain amount of R.A. Typically, the observed sources have declinations $\delta \ll +66.7$, which means that they culminate well above the H.E.S.S. horizon ($\phi \approx 23^{\circ}.3$). In those cases, the run duration is 25–28 min, and the shift in R.A. is usually $\pm 30'$. The edges of the ON and OFF FoV are separated by 2.5° so that there is no overlap. The final product of the so-called ON/OFF observation is a pair of files that contains data separated both spatially and temporally. The OFF portion contains the background and the ON portion both the background and, possibly, the signal. The method relies on the small angular and temporal difference between the ON and OFF observations and assumes that the background levels should be similar in both. To ensure a better similarity, the time difference between the duration of a run and the shift in R.A. is held up in order for the two observations to have identical azimuth and altitude profiles (Fig. 4.51). Moreover, because both background and on-source data are recorded by the same part of the camera, the hardware differences across the camera do not introduce a bias to either observation.

However, instruments with large FoV, like H.E.S.S., have the capability to observe a source using the wobble mode [368]. During a wobble observation, the telescope is pointing at the part of the sky which is ahead or behind the source by 0.5° declination. Then, the on-source portion of the camera is defined around the point which is half a degree away from the camera centre, along the declination

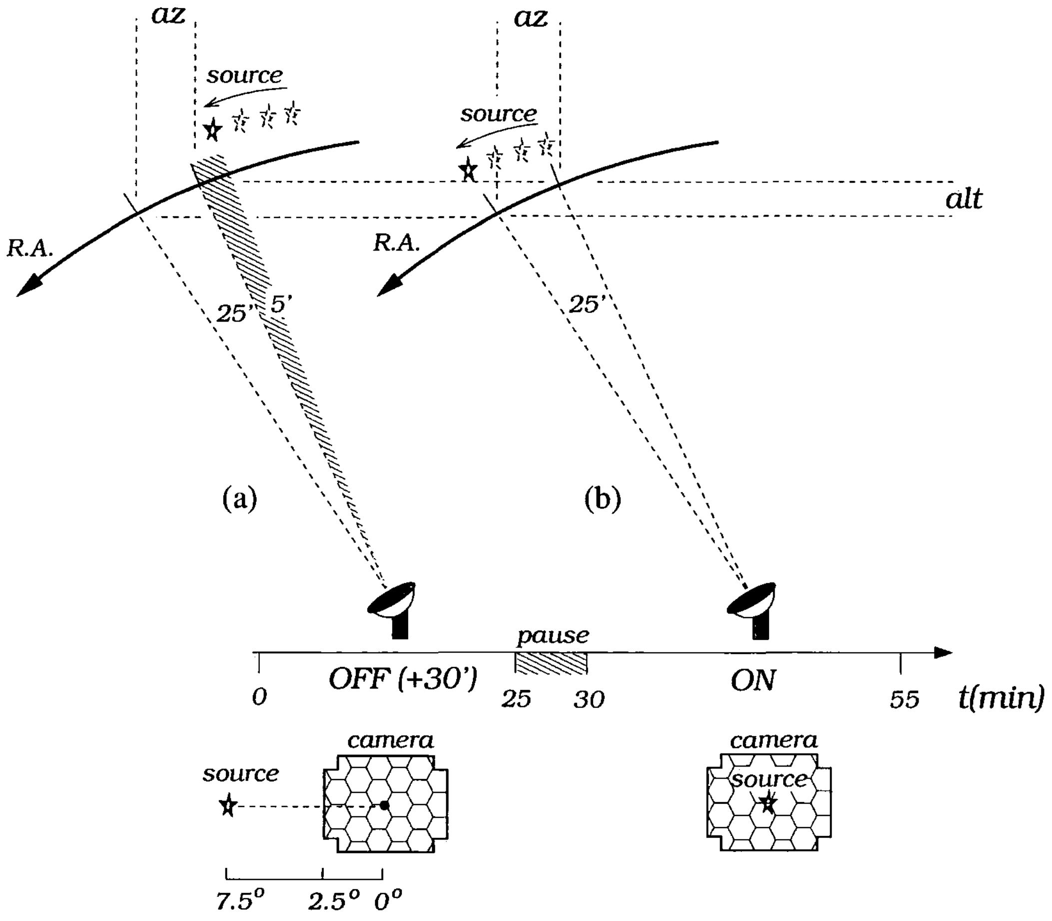


Figure 4.51: Example of an ON/OFF observation. (a) First, the telescope points at 30′ in R.A. ahead of the source, in order to record 25 min of background. During this OFF observation the source is kept 5° away from the edge of the camera. After a pause of 5 min the telescope points at the source, which is at the exact same position where the OFF observation started 30 min earlier. (b) Then, the ON observation tracks the source along the same path as before, for another 25 min. This time, the source is kept at the centre of the FoV. The synchronisation between the time delay and the angular difference between the ON and the OFF observations results in identical altitude and azimuth profiles.

axis. The OFF observation is taken from the symmetrically opposite location on the camera, always with respect to the camera centre. The extent of the ON and OFF regions is defined by the appropriate *alpha* or θ^2 cut, depending on the number of telescopes (see section 4.2.6). Finally, the ON and OFF areas on the camera are swapped around in successive runs, in order to reduce hardware biases associated with the hardware differences across the camera (Fig. 4.52a).

The clear advantage of wobble-mode observations relative to ON/OFF ones is that the observation time for a specific object is doubled, since there is no need for a separate off-source observation. Furthermore, the background estimation is more reliable, because it comes from an area near to the source, which is simultaneously sampled with the latter. Last but not least, the exposure ([area]×[exposure time]) of both ON and OFF data sets is identical, since they come from the same run, so there is no need for normalisation.

Once H.E.S.S. Phase I was completed, it became possible to observe a source in stereoscopic wobble mode. The principle is the same, and the increased angular resolution of the system helps towards a better localisation of the on-source area, via a tighter θ^2 cut. Moreover, it is possible to further reduce the systematic errors in the background estimation, and at the same time avoid multiple observations with alternating offsets, by defining a background area of different shape to the on-source one. The OFF portion of the observation is taken from an annulus of customisable width and radius (typically $r = 0.5°$), which is centred on the source position. That way, there is no preferred direction from which the background is sampled, and hence this eliminates the need for additional observations. Since the ON and OFF areas are unequal, a normalisation factor, $A_{\mathrm{ON}}/A_{\mathrm{OFF}}$, is required prior to any direct comparison between the ON and OFF event rates. Furthermore, the hardware differences between the different parts of the OFF area have to be taken into account, because the sensitivity variations between the different pixels bias the data rate. This mode of background estimation is called *ring* background model [369] (see Fig. 4.52b). It should be mentioned that in addition to those three, there are several other ways to estimate the background, but they will not be discussed here. The reader is directed to [368] for a more detailed presentation of the existing observation modes.

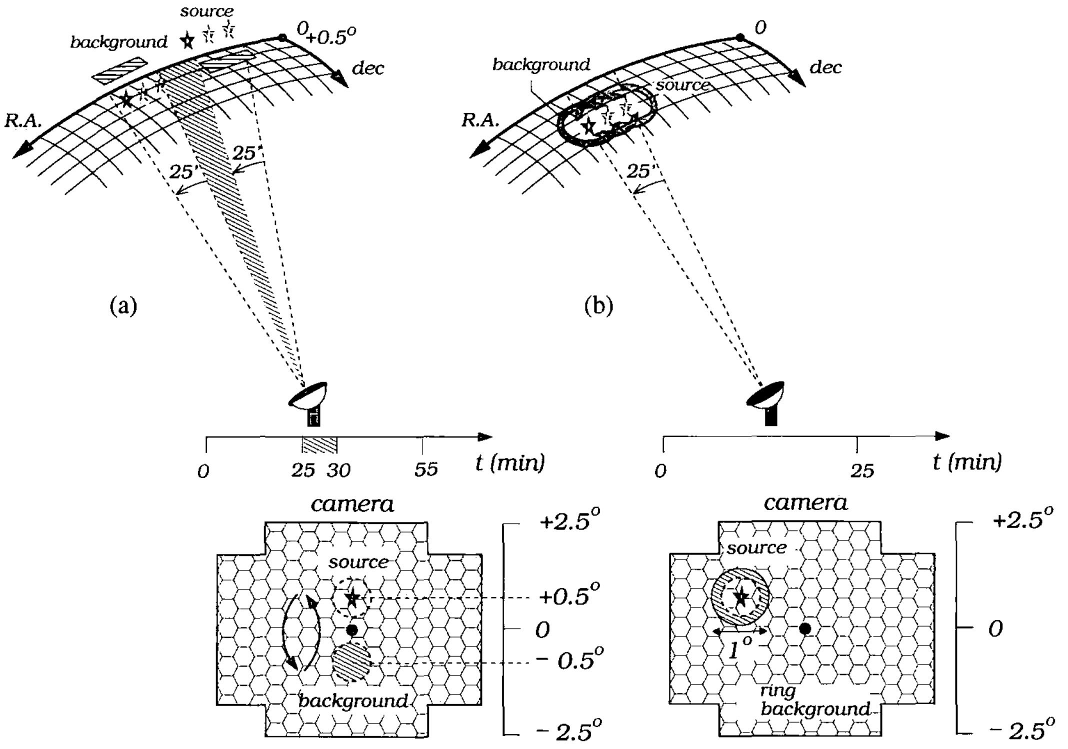


Figure 4.52: Two popular wobble observation modes. (a) In the wobble mode with simple background estimation, the source is initially offset with respect to the centre of the FoV by ±0.5° declination, and the background is measured from the symmetrically opposite area with declination offset of ∓0.5°. Then, a successive observation run is made, with the source and the background offset swapped around. This way the background is sampled from either side of the source position, which helps to minimise systematic errors. (b) In the wobble mode with Ring background estimation, the background is sampled from an annulus of radius 0.5° and of customisable width, centred on the source. The symmetry of the background region means that any directional bias associated with background sampling from a specific area in the sky is removed. Hence, multiple observations like in wobble mode are not required. However, the extended nature of the background area means that hardware differences along the annulus have to be accounted for.